REMARKS

Applicants have amended the specification to recognize the Trademark ATCC™.

Applicants have cancelled Claims 1-3 without prejudice to, or disclaimer of, the subject matter contained therein. Applicants maintain that the cancellation of a claim makes no admission as to its patentability and reserve the right to pursue the subject matter of the cancelled claim in this or any other patent application.

Applicants have amended Claim 4 to be in independent form, and have amended Claims 5 and 12 to depend from Claim 4. Claim 13 is amended to replace the term "epitope tag" with the term "tag polypeptide." New Claims 14-17 have been added. Thus, Claims 4-8 and 11-17 are presented for further examination.

Applicants submit that no new matter was added by the amendments, and that support for the amendments can be found throughout the specification. Support for the amendments to Claim 13 can be found, for example, at paragraph [0229]. Support for new Claims 14-17 can be found, for example, in the claims as originally filed and paragraphs [0336], [0362], [407], and Example 18 starting at paragraph [0529].

Applicants thank the Examiner for the review of the instant application, and withdrawing the objections to the claims, title and specification. Applicants acknowledge the Examiner's withdrawal of the rejection of Claims 1-13 under 35 U.S.C. § 112, second paragraph, as being indefinite. The rejections of the presently pending claims are respectfully traversed.

Information Disclosure Statement

The Examiner has stated that the new patent listed in the information disclosure statement filed October 4, 2004 (mailed by Applicants on September 30, 2004), has not been fully considered since Applicants did not submit the requisite fee. Applicants submit that the October 4, 2004 form PTO-1449 did not disclose a "new" patent, but rather corrected a typographical error in the form PTO-1449 submitted on September 6, 2002, as explained in the September 30, 2004 Office Action Response. A copy of the correct patent was submitted with the original form PTO-1449 in September, 2002. Thus, Applicants are not asking the PTO to consider a "new" patent. As such, Applicants believe that no fee is due, and respectfully request consideration of the patent on file.

Specification

Trademarks:

The Examiner has noted that while the use of trademarks is permissible in patent applications, the trademarks should be capitalized and accompanied by the generic terminology. The Examiner has pointed out that "ATCC" is a trademark, and needs to be recognized wherever it appears. Applicants have amended the specification accordingly.

Priority

The Examiner maintains that Applicants are not entitled to a priority date earlier than the May 2, 2002 filing date of the instant application. Specifically, the Examiner argues that the provisional application filed June 24, 1998 only describes SEQ ID NO: 56 and fails to show utility and enablement for the presently claimed polypeptides. Similarly, the Examiner asserts that Applicants are not entitled to the priority filing date of August 24, 2000, in which the data presented in Example 18 were first disclosed. According to the Examiner, the PCT application does not comply with 35 U.S.C. § 120 written description, utility and enablement.

For the reasons set forth below, Applicants submit that the presently claimed invention meets the requirements set forth in 35 U.S.C. §§ 101 and 112, first paragraph. Accordingly, Applicants are entitled to a priority date of at least August 24, 2000.

Rejection Under 35 U.S.C. §101 – Utility

The Examiner maintains the rejection of Claims 1-8 and 11-13 under 35 U.S.C. § 101 as lacking utility for the reasons set forth on pages 5-7 of the previous Office Action. The Examiner asserts that further research is necessary to confirm a "real world" utility for the claimed polypeptides, and thus they lack substantial utility. The Examiner alleges that Applicants have provided a single analysis of [a] nucleic acid without any relative range for basing a utility of under-expression for the claimed polypeptides, and that no levels (relative or absolute) are disclosed. The Examiner asserts that applicants have provided "no guidance on how to use [the data presented in Example 18]." Office Action at 5.

Next, the Examiner asserts that the functional language "wherein the nucleic acid encodes a polypeptide that is more highly expressed in normal skin" does not confer utility on the claimed polypeptides because there is no data regarding protein expression in melanoma and normal skin. The Examiner maintains that Applicants cannot rely on a correlation between mRNA levels and increased polypeptide levels to establish that PRO1027 is more highly expressed in normal skin than

in melanoma. According to the Examiner "[t]he art clearly establishes that DNA copy number, mRNA levels and protein levels are not inexorably related, in that an increase in one necessarily leads to an increase in all of them. Transcription levels (mRNA) do not correlate with polypeptide levels." Id. (emphasis added). The Examiner relies upon six references to support the proposition that "one skilled in the art would not associate DNA copy number, mRNA and protein levels and necessarily reflecting each other" [sic]: (1) Haynes et al. (1998); (2) an excerpt from Lewin, Genes VI (1997); (3) Gökman-Polar et al. (2001); (4) Pennica et al. (1998); (5) Hu et al. (2003); and (6) Konopka et al (1986).

The Examiner also alleges that the data in the specification is not scientifically significant and does not teach the expression level of the claimed protein. The Examiner alleges that the data in Example 18 is a preliminary observation that requires "substantive experimentation to ascertain the veracity thereof." Office Action at 7. Further, the Examiner states that the visual quantification using ethidium bromide staining of PCR products on gels is "not persuasive" and is "highly subjective." Office Action at 8. Further, according to the Examiner, without knowing the range of variation of transcript levels, there is insufficient guidance as to the substantial utility of the claimed invention. Office Action at 9.

The Examiner rejects Applicants assertion that regardless of whether the polypeptide is over-expressed in melanoma, the polypeptide has utility based on the differential expression of the encoding mRNA. Specifically, the Examiner argues that unlike Her2/Neu, no mutation or translocation of the gene encoding PRO1027 has been associated with tumor formation or the development of cancer.

Utility – Legal Standard

According to the Utility Examination Guidelines ("Utility Guidelines"), 66 Fed. Reg. 1092 (2001) an invention complies with the utility requirement of 35 U.S.C. § 101, if it has at least one asserted "specific, substantial, and credible utility" or a "well-established utility."

Under the Utility Guidelines, a utility is "specific" when it is particular to the subject matter claimed. For example, it is generally not enough to state that a nucleic acid is useful as a diagnostic tool without also identifying the condition that is to be diagnosed.

The requirement of "substantial utility" defines a "real world" use, and derives from the Supreme Court's holding in *Brenner v. Manson*, 383 U.S. 519, 534 (1966) stating that "The basic *quid pro quo* contemplated by the Constitution and the Congress for granting a patent monopoly

is the benefit derived by the public from an invention with substantial utility." In explaining the "substantial utility" standard, M.P.E.P. § 2107.01 cautions, however, that Office personnel must be careful not to interpret the phrase "immediate benefit to the public" or similar formulations used in certain court decisions to mean that products or services based on the claimed invention must be "currently available" to the public in order to satisfy the utility requirement. "Rather, *any reasonable use that an applicant has identified for the invention that can be viewed as providing a public benefit should be accepted as sufficient,* at least with regard to defining a 'substantial' utility." (M.P.E.P. § 2107.01, emphasis added).

Indeed, the Guidelines for Examination of Applications for Compliance With the Utility Requirement, set forth in M.P.E.P. § 2107 II(B)(1) gives the following instruction to patent examiners: "If the applicant has asserted that the claimed invention is useful for any particular practical purpose ... and the assertion would be considered credible by a person of ordinary skill in the art, do not impose a rejection based on lack of utility."

Utility need NOT be Proved to a Statistical Certainty – a Reasonable Correlation between the Evidence and the Asserted Utility is Sufficient

An Applicant's assertion of utility creates a presumption of utility that will be sufficient to satisfy the utility requirement of 35 U.S.C. § 101, "unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope." *In re Langer*, 503 F.2d 1380, 1391, 183 USPQ 288, 297 (CCPA 1974). *See, also In re Jolles*, 628 F.2d 1322, 206 USPQ 885 (CCPA 1980); *In re Irons*, 340 F.2d 974, 144 USPQ 351 (1965); *In re Sichert*, 566 F.2d 1154, 1159, 196 USPQ 209, 212-13 (CCPA 1977). Compliance with 35 U.S.C. § 101 is a question of fact. *Raytheon v. Roper*, 724 F.2d 951, 956, 220 USPQ 592, 596 (Fed. Cir. 1983) cert. denied, 469 US 835 (1984). The evidentiary standard to be used throughout *ex parte* examination in setting forth a rejection is a preponderance of the evidence, or "more likely than not" standard. *In re Oetiker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). This is stated explicitly in the M.P.E.P.:

> [T]he applicant does not have to provide evidence sufficient to establish that an asserted utility is true "beyond a reasonable doubt." **Nor must the applicant provide evidence such that it establishes an asserted utility as a matter of statistical certainty**. Instead, evidence will be sufficient if, considered as a whole, it leads a person of ordinary skill in the art to conclude that the asserted

> utility is more likely than not. M.P.E.P. at § 2107.02, part VII (2004) (underline emphasis in original, bold emphasis added, internal citations omitted).

The PTO has the initial burden to offer evidence "that one of ordinary skill in the art would reasonably doubt the asserted utility." *In re Brana*, 51 F.3d 1560, 1566, 34 U.S.P.Q.2d 1436 (Fed. Cir. 1995). Only then does the burden shift to the Applicant to provide rebuttal evidence. *Id.* As stated in the M.P.E.P., such rebuttal evidence does not need to absolutely prove that the asserted utility is real. Rather, the evidence only needs to be reasonably indicative of the asserted utility.

In *Fujikawa v. Wattanasin*, 93 F.3d 1559, 39 U.S.P.Q.2d 1895 (Fed. Cir. 1996), the Court of Appeals for the Federal Circuit upheld a PTO decision that *in vitro* testing of a novel pharmaceutical compound was sufficient to establish practical utility, stating the following rule:

> [T]esting is often required to establish practical utility. But the test results **need not absolutely prove** that the compound is pharmacologically active. All that is required is that the tests be "*reasonably* indicative of the desired [pharmacological] response." In other words, there must be **a sufficient correlation** between the tests and an asserted pharmacological activity so as to convince those skilled in the art, **to a reasonable probability**, that the novel compound will exhibit the asserted pharmacological behavior." *Fujikawa v. Wattanasin*, 93 F.3d 1559, 1564, 39 U.S.P.Q.2d 1895 (Fed. Cir. 1996) (internal citations omitted, bold emphasis added, italics in original).

While the *Fujikawa* case was in the context of utility for pharmaceutical compounds, the principals stated by the Court are applicable in the instant case where the asserted utility is for a therapeutic and diagnostic use – utility does not have to be established to an absolute certainty, rather, the evidence must convince a person of skill in the art "to a reasonable probability." In addition, the evidence need not be direct, so long as there is a "sufficient correlation" between the tests performed and the asserted utility.

The Court in *Fujikawa* relied in part on its decision in *Cross v. Iizuka*, 753 F.2d 1040, 224 U.S.P.Q. 739 (Fed. Cir. 1985). In *Cross*, the Appellant argued that basic *in vitro* tests conducted in cellular fractions did not establish a practical utility for the claimed compounds. Appellant argued that more sophisticated *in vitro* tests using intact cells, or *in vivo* tests, were necessary to establish a practical utility. The Court in *Cross* rejected this argument, instead favoring the argument of the Appellee:

> [*I*]*n vitro* results...are generally predictive of *in vivo* test results, i.e., there is a **reasonable correlation** therebetween. Were this not so, the testing procedures of the pharmaceutical industry would not be as they are. [Appellee] has not urged, and rightly so, that there is an invariable exact correlation between *in vitro* test results and *in vivo* test results. Rather, [Appellee's] position is that successful *in vitro* testing for a particular pharmacological activity establishes a **significant probability** that *in vivo* testing for this particular pharmacological activity will be successful. *Cross v. Iizuka*, 753 F.2d 1040, 1050, 224 U.S.P.Q. 739 (Fed. Cir. 1985) (emphasis added).

The *Cross* case is very similar to the present case. Like *in vitro* testing in the pharmaceutical industry, those of skill in the field of biotechnology rely on the reasonable correlation that exists between gene expression and protein expression (see below). Were there no reasonable correlation between the two, the techniques that measure gene levels such as microarray analysis, differential display, and quantitative PCR would not be so widely used by those in the art. As in *Cross*, Applicants here do not argue that there is "an invariable exact correlation" between gene expression and protein expression. Instead, Applicants' position detailed below is that a measured change in gene expression in cancer cells establishes a "significant probability" that the expression of the encoded polypeptide in cancer will also be changed based on "a reasonable correlation therebetween."

Taken together, these cases illustrate that the legal standard for demonstrating utility is a relatively low hurdle. An Applicant need only provide evidence such that it is **more likely than not that a person of skill in the art would be convinced, to a reasonable probability, that the asserted utility is true**. The evidence need not be direct evidence, so long as there is a reasonable correlation between the evidence and the asserted utility. The Applicant **does not need to provide evidence such that it establishes an asserted utility as a matter of statistical certainty.**

Even assuming that the Examiner satisfied the initial burden of offering evidence that one of ordinary skill in the art would reasonably doubt the truth of the asserted utility, Applicants assert that they have met their burden of providing rebuttal evidence such that it is more likely than not those skilled in the art, to a reasonable probability, would believe that the claimed invention is useful as a diagnostic tool for cancer.

Substantial Utility

Summary of Applicants' Arguments and the Examiner's Response

In an attempt to clarify Applicants' argument, Applicants offer a summary of their argument and the disputed issues involved. Applicants assert that the claimed polypeptides have utility as diagnostic tools for cancer, particularly skin cancer. Applicants are not asserting that the claimed polypeptides necessarily provide a definitive diagnosis of cancer, but rather that they are useful, alone or in combination with other diagnostic tools to assist in the diagnosis of skin cancer. Applicants' asserted utility rests on the following argument:

1. Applicants have provided reliable evidence that mRNA for the PRO1027 polypeptide is more highly expressed in melanoma compared to normal skin;

2. Applicants assert that it is well-established in the art that a change in the level of mRNA for a particular protein, e.g. an increase, generally leads to a corresponding change in the level of the encoded protein, e.g. an increase;

3. Given Applicants' evidence that the level of mRNA for the PRO1027 polypeptide is increased in melanoma compared to normal skin, it is likely that the PRO1027 polypeptide is differentially expressed in melanoma and is therefore useful as a diagnostic tool to distinguish melanoma from normal skin.

Applicants understand the Examiner to be making several arguments in response to Applicants' asserted utility:

1. The Examiner challenges the reliability of the evidence reported in Example 18, and states that it does not provide information sufficient to allow a skilled artisan to diagnose any disease;

2. The Examiner cites Haynes et al., Lewin (Genes VI), Gökman-Polar, Pennica, Hu, and Konopka et al. to support the position that the mRNA expression data does not provide utility to the claimed polypeptides because DNA, mRNA and protein levels do not necessarily correlate;

3. The Examiner argues that because no mutation or translocation of PRO1027 has been associated with melanoma, the disclosure is insufficient to meet the utility requirement;

4. The Examiner argues that the specification and the art do not teach how the levels of PRO1027 fit into characterization of tumors, or better determination of therapy.

As detailed below, Applicants submit that the PTO has failed to meet its initial burden to offer evidence "that one of ordinary skill in the art would reasonably doubt the asserted utility."

In re Brana, 51 F.3d 1560, 1566, 34 U.S.P.Q.2d 1436 (Fed. Cir. 1995). First, the Examiner has failed to offer any evidence to support its rejection of the data in Example 18 and the Declaration of Chris Grimaldi in support of these data. Second, the fact that no mutation or translocation of PRO1027 has been associated with melanoma does not prevent its use as a diagnostic tool for the same. Third, Applicants submit that given the well-established correlation between a change in the level of mRNA with a corresponding change in the levels of the encoded protein, the PRO1027 protein is likely differentially expressed in melanoma compared to normal skin. This provides utility for the PRO1027 and related proteins as cancer diagnostic tools. Finally, even if the PTO has met its initial burden, Applicants have submitted enough rebuttal evidence to establish that it is **more likely than not** that a person of skill in the art would be convinced, **to a reasonable probability,** that the asserted utility is true. As stated above, Applicants' evidence need not be direct evidence, so long as there is a reasonable correlation between the evidence and the asserted utility. **The standard is not statistical or absolute certainty.**

Applicants have established that the Gene Encoding the PRO1027 Polypeptide is Differentially Expressed in Certain Cancers compared to Normal Tissue

Applicants first address the Examiner's argument that the evidence of differential expression of the gene encoding the PRO1027 polypeptide melanoma is insufficient to support a substantial utility.

The gene expression data in the specification, Example 18, shows that the mRNA associated with protein PRO1027 was more highly expressed in melanoma compared to normal skin tissue. Gene expression was analyzed using standard semi-quantitative PCR amplification reactions of cDNA libraries isolated from different melanoma and normal skin tissue samples. Identification of the differential expression of the PRO1027 polypeptide-encoding gene in tumor tissue compared to the corresponding normal tissue renders the molecule useful as a diagnostic tool for the determination of the presence or absence of tumor. In support, Applicants previously submitted as Exhibit 1 a first Declaration of J. Christopher Grimaldi, an expert in the field of cancer biology. This declaration explains the importance of the data in Example 18, and how differential gene and protein expression studies are used to differentiate between normal and tumor tissue (see Declaration, paragraph 7).

In paragraph 5 of his declaration, Mr. Grimaldi states that the gene expression studies reported in Example 18 of the instant application were made from pooled samples of normal and of tumor tissues. Mr. Grimaldi explains that:

> The DNA libraries used in the gene expression studies were made from pooled samples of normal and of tumor tissues. *Data from pooled samples is more likely to be accurate than data obtained from a sample from a single individual.* That is, the detection of variations in gene expression is likely to represent a more generally relevant condition when pooled samples from normal tissues are compared with pooled samples from tumors in the same tissue type. (Paragraph 5) (emphasis added).

In paragraphs 6 and 7, Mr. Grimaldi explains that the semi-quantitative analysis employed to generate the data of Example 18 is sufficient to determine if a gene is over- or under-expressed in tumor cells compared to corresponding normal tissue. He states that any visually detectable difference seen between two samples is indicative of at least a two-fold difference in cDNA between the tumor tissue and the counterpart normal tissue. Thus, the results of Example 18 reflect at least a two-fold difference in PRO1027 mRNA levels between normal and tumor samples. He also states that the results of the gene expression studies indicate that the genes of interest "can be used to differentiate tumor from normal," thus establishing their reliability. He explains that "The precise levels of gene expression are irrelevant; what matters is that there is a relative difference in expression between normal tissue and tumor tissue." (Paragraph 7). Thus, since it is the relative level of expression between normal tissue and suspected cancerous tissue that is important, the precise level of expression in normal tissue is irrelevant. Likewise, there is no need for quantitative data to compare the level of expression in normal and tumor tissue. As Mr. Grimaldi states, "If a difference is detected, this indicates that the gene and its corresponding polypeptide and antibodies against the polypeptide are useful for diagnostic purposes, to screen samples to differentiate between normal and tumor."

Nevertheless, the Examiner maintains that the Grimaldi Declaration is insufficient to overcome the rejection of Claims 1-8 and 11-13. The Examiner argues that Example 18 is insufficient because it does not teach how high the expression level is, the expression range for normal and tumor tissue, whether the results are statistically significant, or the number of samples that were used. The Examiner also argues that one cannot determine if the observed increase in PRO1027 "nucleic acid" is due to a mutation, copy number differences (presumably DNA), or an increase in transcription rates. The Examiner concludes that the disclosure would

not enable one of skill in the art to differentiate amongst expression levels to diagnose any disease. The Examiner also alleges that the Grimaldi declaration is unpersuasive as it does not address the PRO1027 polypeptide.

Applicants submit that the declaration of Mr. Grimaldi is based on personal knowledge of the relevant facts at issue. Mr. Grimaldi is an expert in the field and conducted or supervised the experiments at issue. Applicants remind the Examiner that "[o]ffice personnel must accept an opinion from a qualified expert that is based upon relevant facts whose accuracy is not being questioned." PTO Utility Examination Guidelines (2001) (emphasis added). In addition, declarations relating to issues of fact should not be summarily dismissed as "opinions" without an adequate explanation of how the declaration fails to rebut the Examiner's position. *In re Alton* 76 F.3d 1168 (Fed. Cir. 1996). The Examiner has not supplied any reasons or evidence to question the accuracy of the facts upon which Mr. Grimaldi based his opinion. Mr. Grimaldi has personal knowledge of the relevant facts, has based his opinion on those facts, and the Examiner has offered no reason or evidence to reject either the underlying facts or his opinion, other than summarily stating that visual detection is highly subjective. Applicants note that the statement that any visually detectable difference between two samples is indicative of at least a two-fold difference in cDNA is a statement of scientific fact. As discussed above, a declaration that is filed to address a question of fact must be considered by the Examiner and if the Examiner maintains his rejection, he must be able to explain why the declaration fails. *In re Alton* 76 F.3d 1168 (Fed. Cir.1996). In addition, Applicants provide herewith as Exhibit 1 a copy of page 122 of the 2002-2003 New England Biolabs catalog. Exhibit 1 shows DNA size markers of differing lengths run on an agarose gel. The column on the left provides the mass of each marker in nanograms and the column on the right provides the length of the marker. It is apparent that the band intensity of markers having mass differences of two fold are readily distinguishable by eye (See for example, the difference in band intensities of the 0.1kb fragment present at 61ng and the 0.5kb marker present at 124ng). Accordingly, Applicants maintain that the procedures used to detect differences in expression levels were sufficiently sensitive to detect two-fold differences.

In light of the above, the Examiner should accept Mr. Grimaldi's opinion with regard to his statement that "any visually detectable difference seen between two samples is indicative of at least a two-fold difference in cDNA between the tumor tissue and the counterpart normal tissue" and that the genes of interest "can be used to differentiate tumor from normal." Together,

these statements establish that there is at least a two-fold difference in expression, and that the results are reliable enough that they can be used to distinguish tumor from normal tissue. Furthermore, Applicants submit that the first Grimaldi declaration addresses PRO1027, as well as the other PRO polypeptides listed in Example 18. Paragraphs 4-7 relate to the data reported in Example 18 of the specification. Example 18 teaches that the nucleic acids encoding various PRO polypeptides, including PRO1027, are differentially expressed in normal tissue as compared to cancer tissue samples. As such, the first Grimaldi declaration addresses the claimed invention.

In conclusion, Applicants submit that the evidence reported in Example 18, combined with the first Grimaldi Declaration, establish that there is at least a two-fold difference in PRO1027 cDNA between melanoma and normal skin tissue. Therefore, it follows that expression levels of the PRO1027 gene can be used to distinguish melanoma from normal skin. The Examiner has not offered any significant arguments or evidence to the contrary.

As Applicants explain below, it is more likely than not that the PRO1027 polypeptide is also differentially expressed in melanoma, and can therefore be used to distinguish melanoma from normal skin. This provides utility for the claimed polypeptides.

Applicants have established that the Accepted Understanding in the Art is that there is a Direct Correlation between mRNA Levels and the Level of Expression of the Encoded Protein

Applicants next turn to the second portion of their argument in support of their asserted utility – that it is well-established in the art that a change in the level of mRNA for a particular protein generally leads to a corresponding change in the level of the encoded protein; given Applicants' evidence of differential expression of the mRNA for the PRO1027 polypeptide in melanoma compared to normal skin, it is more likely than not that the PRO1027 polypeptide is differentially expressed; and proteins differentially expressed in certain tumors have utility as diagnostic tools.

In support of the assertion that changes in mRNA are positively correlated to changes in protein levels, Applicants previously submitted a copy of a second Declaration by J. Christopher Grimaldi, an expert in the field of cancer biology (previously attached as Exhibit 2). This declaration also addresses the data in Example 18, and therefore specifically addresses Example 18. As stated in paragraph 5 of the declaration, "Those who work in this field are well aware that in the vast majority of cases, when a gene is over-expressed...the gene product or polypeptide

will also be over-expressed.... This same principal applies to gene under-expression." Further, "the detection of increased mRNA expression is expected to result in increased polypeptide expression, and the detection of decreased mRNA expression is expected to result in decreased polypeptide expression. The detection of increased or decreased polypeptide expression can be used for cancer diagnosis and treatment." The references cited in the declaration and submitted herewith support this statement.

Applicants also previously submitted a copy of the declaration of Paul Polakis, Ph.D. (previously attached as Exhibit 3), an expert in the field of cancer biology. As stated in paragraph 6 of his declaration:

> Based on my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4 and 5 above [showing a positive correlation between mRNA levels and encoded protein levels in the vast majority of cases] and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor cell relative to a normal cell typically correlates to a similar increase in abundance of the encoded protein in the tumor cell relative to the normal cell. In fact, *it remains a central dogma in molecular biology that increased mRNA levels are predictive of corresponding increased levels of the encoded protein.* (Emphasis added).

Dr. Polakis acknowledges that there are published cases where such a correlation does not exist, but states that it is his opinion, based on over 20 years of scientific research, that "such reports are exceptions to the commonly understood general rule that increased mRNA levels are predictive of corresponding increased levels of the encoded protein." (Polakis Declaration, paragraph 6).

The statements of Grimaldi and Polakis are supported by the teachings in Molecular Biology of the Cell, a leading textbook in the field (Bruce Alberts, *et al.*, Molecular Biology of the Cell (3rd ed. 1994) (submitted herewith as Exhibit 2) and (4th ed. 2002) (submitted herewith as Exhibit 3). Figure 9-2 of Exhibit 2 shows the steps at which eukaryotic gene expression can be controlled. The first step depicted is transcriptional control. Exhibit 2 provides that "[f]or most genes transcriptional controls are paramount. This makes sense because, of all the possible control points illustrated in Figure 9-2, only transcriptional control ensures that no superfluous intermediates are synthesized." Exhibit 2 at 403 (emphasis added). In addition, the text states that "Although controls on the initiation of gene transcription are the predominant form of regulation for most genes, other controls can act later in the pathway from RNA to protein to

modulate the amount of gene product that is made." Exhibit 2 at 453 (emphasis added). Thus, as established in Exhibit 2, the predominant mechanism for regulating the amount of protein produced is by regulating transcription initiation.

In Exhibit 3, Figure 6-3 on page 302 illustrates the basic principle that there is a correlation between increased gene expression and increased protein expression. The accompanying text states that "a cell can change (or regulate) the expression of each of its genes according to the needs of the moment – *most obviously by controlling the production of its mRNA*." Exhibit 3 at 302 (emphasis added). Similarly, Figure 6-90 on page 364 of Exhibit 3 illustrates the path from gene to protein. The accompanying text states that while potentially each step can be regulated by the cell, "the initiation of transcription is the most common point for a cell to regulate the expression of each of its genes." Exhibit 3 at 364 (emphasis added). This point is repeated on page 379, where the authors state that of all the possible points for regulating protein expression, "[f]or most genes transcriptional controls are paramount." Exhibit 3 at 379 (emphasis added).

Further support for Applicants' position can be found in the textbook, Genes VI, (Benjamin Lewin, Genes VI (1997)) (submitted herewith as Exhibit 4) which states "having acknowledged that control of gene expression can occur at multiple stages, and that production of RNA cannot inevitably be equated with production of protein, it is clear that the overwhelming majority of regulatory events occur at the initiation of transcription." *Genes VI* at 847-848 (emphasis added).

Additional support is also found in Zhigang *et al.*, World Journal of Surgical Oncology 2:13, 2004, submitted herewith as Exhibit 5. Zhigang studied the expression of prostate stem cell antigen (PSCA) protein and mRNA to validate it as a potential molecular target for diagnosis and treatment of human prostate cancer. The data showed "a high degree of correlation between PSCA protein and mRNA expression." Exhibit 5 at 6. Of the samples tested, 81 out of 87 showed a high degree of correlation between mRNA expression and protein expression. The authors conclude that "it is demonstrated that PSCA protein and mRNA overexpressed in human prostate cancer, and that the increased protein level of PSCA was resulted from the upregulated transcription of its mRNA." Exhibit 5 at 11. Even though the correlation between mRNA expression and protein expression occurred in 93% of the samples tested, not 100%, the authors

state that "PSCA may be a promising molecular marker for the clinical prognosis of human Pca and a valuable target for diagnosis and therapy of this tumor." *Id.*

Further, Meric *et al.*, Molecular Cancer Therapeutics, vol. 1, 971-979 (2002), submitted herewith as Exhibit 6, states the following:

> The **fundamental principle** of molecular therapeutics in cancer is to exploit the differences in gene expression between cancer cells and normal cells...[M]ost efforts have concentrated on identifying differences in gene expression at the level of mRNA, which can be attributable to either DNA amplification or to differences in transcription. Meric *et al.* at 971 (emphasis added).

Those of skill in the art would not be focusing on differences in gene expression between cancer cells and normal cells if there were no correlation between gene expression and protein expression.

Together, the declarations of Grimaldi and Polakis, the accompanying references, and the excerpts and references provided above all establish that the accepted understanding in the art is that there is a reasonable correlation between changes in gene expression and the level of the encoded protein.

In response to the second Grimaldi Declaration, the Examiner states that unlike Her2/Neu, PRO1027 "has not been associated with tumor formation or the development of cancer." Office Action at 10. The Examiner also maintains that regarding t(5;14) discussed in the second Grimaldi declaration, "[n]o mutation or translocation of PRO1027 is known or established to occur." Id. The Examiner concludes that in the absence of any of the above information, the disclosure is insufficient to meet the utility requirement. Applicants submit that a lack of known role for PRO1027 in cancer does not prevent its use as a diagnostic tool for cancer. The fact that there is no known translocation or mutation of PRO1027 is irrelevant to whether its differential expression can be used to assist in diagnosis of cancer – one does not need to know why PRO1027 is differentially expressed, or what the consequence of the differential expression is, in order to exploit the differential expression to distinguish tumor from normal tissue. In fact the Revised Interim Utility Guidelines promulgated by the PTO recognize that proteins which are differentially expressed in cancer have utility. (*See* the caveat in Example 12 which state that the utility requirement is satisfied where a protein is expressed in melanoma cells but not on normal skin and antibodies against the protein can be used to diagnose cancer.)

In addition, while Applicants appreciate that actions taken in other applications are not binding on the PTO with respect to the present application, Applicants note that the PTO has issued several patents claiming differentially expressed polypeptides. (*See, e.g.,* U.S. Patent No. 6,414,117 and U.S. Patent No. 6,124,433, attached hereto as Exhibits 7 and 8.) Because it is more likely than not that the PRO1027 protein is also differentially expressed in certain tumors, the protein can be used as a diagnostic tool for cancer.

Thus, the Examiner's rejection of the second Grimaldi Declaration and Polakis Declaration because they are viewed as insufficient is misplaced. Accordingly, Applicants submit that they have offered sufficient evidence to establish that it is more likely than not that one of skill in the art would believe that because the PRO1027 mRNA is more highly expressed in melanoma tumor compared to normal skin tissue, the PRO1027 polypeptide will have the same expression pattern. This differential expression of the PRO1027 and related polypeptides make them useful as diagnostic tools for melanoma.

The Arguments made by the PTO are Not Sufficient to satisfy the PTO's Initial Burden of Offering Evidence "that one of ordinary skill in the art would reasonably doubt the asserted utility"

As stated above, an Applicant's assertion of utility creates a presumption of utility that will be sufficient to satisfy the utility requirement of 35 U.S.C. § 101, "unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope." *In re Langer*, 503 F.2d 1380, 1391, 183 USPQ 288, 297 (CCPA 1974). The evidentiary standard to be used throughout *ex parte* examination in setting forth a rejection is a preponderance of the evidence, or "more likely than not" standard. *In re Oetiker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). This is stated explicitly in the M.P.E.P.:

> [T]he applicant does not have to provide evidence sufficient to establish that an asserted utility is true "beyond a reasonable doubt." **Nor must the applicant provide evidence such that it establishes an asserted utility as a matter of statistical certainty**. Instead, evidence will be sufficient if, considered as a whole, it leads a person of ordinary skill in the art to conclude that the asserted utility is more likely than not. M.P.E.P. at § 2107.02, part VII (2004) (underline emphasis in original, bold emphasis added, internal citations omitted).

The PTO has the initial burden to offer evidence "that one of ordinary skill in the art would reasonably doubt the asserted utility." *In re Brana*, 51 F.3d 1560, 1566, 34 U.S.P.Q.2d

1436 (Fed. Cir. 1995). Only then does the burden shift to the Applicant to provide rebuttal evidence. *Id.* As stated in the M.P.E.P., such rebuttal evidence does not need to absolutely prove that the asserted utility is real. Rather, the evidence only needs to be reasonably indicative of the asserted utility.

The Examiner has cited six references to establish "that one of ordinary skill in the art would reasonably doubt" that the disclosed polypeptide is differentially expressed in certain tumors and that the claimed polypeptides can be used as diagnostic tools. Applicants address each cited reference in turn.

The Examiner asserts that Haynes et al. (1998, *Electrophoresis*, 19:1862-1871) found "no strong correlation between protein and transcript level" when studying protein expression in *Saccharomyces cerevisiae*. In particular, the Examiner points to the finding that, for some genes, equivalent mRNA levels translated into protein abundances which varied by more than 50-fold. The Examiner also points to the statement that protein levels cannot be accurately predicted from the level of the corresponding mRNA transcript.

Applicants submit that Haynes does not contradict the utility or enablement of the instant claims. Haynes is a review article dealing with the art of proteome analysis. The assertions in Haynes cited by the Examiner were made in an effort to identify shortcomings in the art of mRNA quantification to argue for "proteome analysis to become an essential component in the comprehensive analysis of biological systems." Haynes, p. 1863. Haynes studied 80 selected samples from *Saccharomyces cerevisiae*, and reported "a general trend but no strong correlation between protein and transcript levels (Fig. 1)." Id. However, a cursory inspection of Fig. 1 shows a clear correlation between the mRNA levels and protein levels measured. This correlation is confirmed by an inspection of the full-length research paper from which the data in Fig. 1 were derived, presented herein as Exhibit 9 (Gygi et al., Molecular and Cellular Biology, Mar. 1999, 1720-1730). Gygi states that "there was a general trend of increased protein levels resulting from increased mRNA levels," with a correlation coefficient of 0.935, indicating a strong correlation. Gygi, p. 1726. Moreover, Gygi also states that the correlation is especially strong for highly expressed mRNAs. Id. Considering that Example 18 of the specification shows over-expression of PRO1027 mRNA in rectal tumor and normal lung, Haynes and Gygi actually provide strong evidence in support of a general correlation between mRNA and protein levels, and thus the utility of the claimed PRO1027 polypeptides and antibodies.

The 50-fold variation referred to by the Haynes and cited by the Examiner, does not in any way show the absence of a correlation between mRNA and protein levels, but rather identifies the outer limits of variability in the Authors' experiments. This variability may support the Authors' assertion that the amount of a particular protein cannot accurately predict the particular level of the corresponding mRNA transcript, but it does not suggest an absence of a general correlation between mRNA and protein levels. Moreover, Gygi states that the high degree of variability seen at low levels of mRNA (shown in inset of Fig. 1, Haynes p. 1863) is due to the fact that "the magnitude of the error in the measurement of mRNA levels is inversely proportional to the mRNA levels." Gygi, p. 1727. Considering that PRO1027 mRNA has been shown in Example 18 of the specification to be over-expressed in rectal tumor and normal lung, the variability identified by Haynes is even less applicable to establishing the absence of a correlation between mRNA and protein levels in the instant case.

As stated above, the standard for utility is not absolute certainty, but rather whether one of skill in the art would be more likely than not to believe the asserted utility. Here, the utility of PRO1027 polypeptides and antibodies as diagnostic tools does not require Applicants to show that mRNA levels correlate to protein levels in every case, but rather only that the correlation exists more often than not. The data presented in Haynes is not inconsistent with or contradictory to the utility or enablement of the instant claims. To the contrary, the data clearly show a general correlation between protein levels and mRNA levels, and thus support Applicants' assertion that such a general correlation exists.

Even if Haynes supported the Examiner's argument, which it does not, one contrary example does not establish that one of skill in the art would find it is more likely than not there is no general correlation between mRNA level and protein levels. In fact, the working hypothesis among those skilled in the art, as illustrated by the evidence presented above by Applicants, is that there is a direct correlation between mRNA levels and protein levels. This is further supported by the statement in Haynes that "interpretations of quantitative mRNA expression profiles frequently implicitly or explicitly assume that for specific genes the transcript levels are indicative of the levels of protein expression." See, Haynes, p. 1863, first full paragraph. Haynes does not suggest there is no correlation between mRNA and protein levels, but rather points to what the authors believe are shortcomings of using mRNA quantification to predict protein levels; specifically, that mRNA levels may not accurately predict protein levels *in each*

particular instance. Considering the more likely than not standard for utility, Haynes' identification of reasons why proteomic analysis may be preferable in some cases does not contradict Applicants' evidence that there is a general correlation between mRNA and protein levels.

The Examiner next offers the statement in Genes VI that, "having acknowledged that control of gene expression can occur at multiple stages, and that production of RNA cannot inevitably be equated with production of protein, it is clear that the overwhelming majority of regulatory events occur at the initiation of transcription," as evidence that mRNA and protein levels do not correlate. (Genes VI, Benjamin Lewin, 1997, Chapter 29 – Regulation of Transcription, 1st page). The Examiner focuses on the statement that "production of RNA cannot inevitably be equated with production of protein." However, the Examiner ignores the statement that, "it is clear that the **overwhelming majority** of regulatory events occur at the initiation of transcription." Further, Lewin goes on to state that transcription of a gene "is a major control point: probably **it is the most common level of regulation**." Id., emphasis added. This reference provides additional support for Applicants' position that that the accepted understanding in the art is that there is a *reasonable* correlation between gene expression and the level of the encoded protein. Applicants do not assert that protein expression is never regulated at the translation level (i.e., post-transcriptionally), but rather that, more often than not, regulation of protein expression occurs at the transcriptional level, making the level of mRNA a general indicator of the level of protein. Far from rebutting this assertion, Lewin in fact supports it by stating that regulation occurs at the transcriptional stage in the overwhelming majority of cases, and that transcription is the most common level of regulation.

The Examiner also cites to a reference authored by Gökmen-Polar et al. (2001, *Cancer Research*, 61: 1375-1381) for the proposition that there is not a necessary correlation between increased mRNA levels and increased protein levels. Gökmen-Polar et al. analyzed the mRNA and protein levels of PKCα, PKCβI and PKCβII in normal versus chemically induced preneoplastic colon tissue. The authors report that "PKCβI and PKCα mRNA levels in [the preneoplastic cells] were decreased to a level commensurate with the decrease in protein expression of these isozymes." Gökmen-Polar et al. at 1378 (emphasis added). The authors also report that level of PKCβII mRNA and protein was increased in the preneoplastic cells, but due to the fact that the increase in protein levels was higher than the increase in mRNA levels, they

conclude that the changes in protein expression levels "could not be explained by transcriptional control alone." Id. at 1380 (emphasis added). In other words, Gökmen-Polar et al. shows a correlation between mRNA and protein levels for each enzyme analyzed, and merely acknowledges the existence of post-transcriptional controls. Thus, rather than providing evidence that "transcription levels (mRNA) do not correlate with polypeptide levels" (Office Action at 5), Gökmen-Polar et al. disclose data that illustrate this correlation.

The Examiner also cites Pennica et al. (1998, *Proc. Nat. Acad. Sci. USA*, 95: 14717-14722) for the proposition that "there is a lack of correlation between gene amplification and protein expression" and that "the asserted correlation between each gene, mRNA and corresponding protein in tumors is unpredictable and well-known to the skilled artisan." Office Action at 6. The evidence in Pennica of a lack of correlation between DNA copy number and mRNA levels says nothing about a lack of correlation between the level of mRNA and the level of protein expression – in fact, Pennica did not even examine protein expression. It is the correlation between mRNA level and the level of protein expression which is at issue here, not the correlation of gene copy number and mRNA levels. The data Applicants report in Example 18 indicate there are more copies of the mRNA encoding PRO1027 polypeptides in rectal tumor compared to normal rectum, and less PRO1027 mRNA in lung tumor compared to normal lung, and thus whether gene amplification leads to increased mRNA is not at issue. Nothing in Pennica is contrary to Applicants' assertion that it is well-established in the art that the level of protein is positively correlated to the level of mRNA. Furthermore, the authors of Pennica expressly qualify the data upon which they based the statement relied upon by the Examiner for the proposition that the field is unpredictable. Specifically, the authors report that the "apparent amplification observed for *WISP-2* may be caused by another gene in [the] amplicon." Pennica at 14722. Applicants submit that Pennica does not establish unpredictability regarding whether DNA copy number and mRNA levels correlate, which Applicants submit is not at issue. Rather, a full reading of Pennica reveals that the authors recognize the limitations of their experiments and expressly cautions readers from drawing the exact conclusion that the Examiner draws from their data. Thus, Applicants submit that Pennica does not establish that it is more likely than not that one skilled in the art would doubt Applicants' asserted utility.

The Examiner cites Hu et al. (2003, *J. Proteome Research*, 2: 405-412) for the proposition that "it is not the norm that increased/decreased gene transcription results in

increased/decreased polypeptide levels." Office Action at 7. Applicants submit that Hu is uninformative regarding the correlation between gene expression level and protein levels. In Hu, the researchers used an automated literature-mining tool to summarize and estimate the relative strengths of all human gene-disease relationships published on Medline. They then generated a microarray expression dataset comparing breast cancer and normal breast tissue. Using their data-mining tool, they looked for a correlation between the strength of the literature association between the gene and breast cancer, and the magnitude of the difference in expression level. They report that for genes displaying a 5-fold change or less in tumors compared to normal, there was no evidence of a correlation between altered gene expression and a *known* role in the disease. *See* Hu at 411. However, among genes with a 10-fold or more change in expression level, there was a strong correlation between expression level and a *published* role in the disease. *Id.* at 412. Importantly, Hu reports that the observed correlation was only found among estrogen receptor-positive tumors, not ER-negative tumors. *Id.*

The general findings of Hu are not surprising – one would expect that genes that have the greatest change in expression in a disease would be the first targets of research, and therefore have the strongest *known* relationship to the disease as measured by the number of reports of a connection in the literature. But this does not mean that genes, and their corresponding proteins, with a lower level of change in expression are not important or cannot be used as molecular markers of the disease. This is demonstrated by the fact that ER-negative tumors did not show a correlation. The correlation reported in Hu only indicates that the greater the change in expression level, the more likely it is that there is a *published* or *known* role for the gene in the disease, as found by their automated literature-mining software. Nowhere in Hu does it say that a lack of correlation in their study means that the genes, and their corresponding proteins, with a less than five-fold change in level of expression in cancer cannot serve as a molecular marker of cancer. Genes with lower levels of change in expression may or may not be the most important genes in causing the disease, but the genes and their corresponding proteins can still show a consistent and measurable change in expression. While such genes and polypeptides may or may not be good targets for further research, they can nonetheless be used as diagnostic tools. Thus, Hu does not refute the Applicants' assertion that the PRO1027 gene, and its corresponding polypeptide, can be used as a cancer diagnostic tool because they are differentially expressed in certain tumors.

The Examiner cites Konopka et al. (1986, *Proc. Nat. Acad. Sci. USA*, 83: 4049-4052) in further support of the assertion that it is not the norm that levels of mRNA correlate with corresponding protein levels. As described for Pennica, the Examiner has confused the relationship between an increase in copy number of a gene and the level of mRNA on the one hand, with the relationship between mRNA expression and levels of the corresponding protein on the other. In particular, the Examiner cites the statement in Konopka that "protein expression is not related to amplification of the abl gene but to variation in the level of the *bcr-abl* mRNA produced from a single Ph1 template." As with Pennica, the results presented in Konopka actually present strong evidence in support of Applicants' position that there is a general understanding in the art that levels of mRNA correlate with levels of the corresponding proteins. Konopka analyzed the expression patterns of a gene associated with certain cancers. The authors show a wide variation in the levels of the protein in various cell types, and find that this variation can be attributed to the levels of the corresponding mRNA in each cell type. See, Konopka, p. 4050. Konopka thus concludes, "these combined data suggest that differential *bcr-abl* mRNA expression from a single gene template is responsible for the variable levels of $P210^{c\text{-}abl}$ [the protein of interest] detected." Id., p. 4051. Thus, far from supporting the Examiner's assertion that it is not the norm that increased transcription leads to increased levels of the corresponding protein, Konopka strongly supports the opposite proposition asserted by Applicants - that the level of mRNA, more often than not, correlates with the level of the corresponding protein.

Thus, the evidence presented by Applicants, including the declarations of Grimaldi and Polakis, the accompanying references, the excerpts from "Molecular Biology of the Cell", and the Zhigang et al. and Meric et al. articles establish that the accepted understanding in the art is that there is a *reasonable* correlation between gene expression and the level of the encoded protein. Further, the references cited by the Examiner as evidence that no such general understanding exists do not in any way support the Examiner's arguments. To the contrary, Haynes, Lewin, Gökmen-Polar, Pennica, Hu and Konopka each provide strong evidence of the correlation asserted by Applicants. Accordingly, Applicants respectfully submit that the totality of the evidence clearly supports the conclusion that one of skill in the art understands that, more likely than not, levels of mRNA directly correlate with levels of corresponding proteins.

Given the lack of support for the Examiner's position, Applicants submit that the Examiner has not met its initial burden of overcoming the presumption that the asserted utility is

sufficient to satisfy the utility requirement. In light of the references cited by Applicants, and further in light of the supporting evidence in the references cited by the Examiner, Applicants submit that the evidence is sufficient to establish that one of skill in the art would be more likely than not to believe that there is a correlation between mRNA levels and polypeptide levels, and thus the claimed polypeptides can be used as diagnostic tools for skin cancer.

Specific Utility

The Asserted Substantial Utilities are Specific to the Claimed Polypeptides

Applicants next address the Examiner's assertion that the asserted utilities are not specific to the claimed polypeptides related to PRO1027. Applicants respectfully disagree.

Specific Utility is defined as utility which is "specific to the subject matter claimed," in contrast to "a general utility that would be applicable to the broad class of the invention." M.P.E.P. § 2107.01 I. Applicants submit that the evidence of differential expression of the PRO1027 gene and polypeptide in certain types of tumor cells, along with the declarations and references discussed above, provide a specific utility for the claimed polypeptides.

As discussed above, there are significant data which show that it is more likely than not that the gene for the PRO1027 polypeptide is more highly melanoma tumor as compared to normal skin. These data are strong evidence that the PRO1027 gene and polypeptide are associated with melanoma. Thus, contrary to the assertions of the Examiner, Applicants submit that they have provided evidence associating the PRO1027 gene and polypeptide with a specific disease. The asserted utility as a diagnostic tool for melanoma is a specific utility – it is not a general utility that would apply to the broad class of polypeptides.

Conclusion

The Examiner has asserted two arguments for why the claimed polypeptides lack utility: (1) that the data reported in Example 18 is not reliable and does not establish a correlation between the differential protein expression and melanoma; (2) that because there is no necessary correlation between gene copy number, mRNA levels and protein levels, the claimed polypeptides are not useful as diagnostic tools;

Applicants have addressed each of these arguments in turn.

First, the Applicants provided a first Declaration of Chris Grimaldi stating that the data in Example 18 are real and significant. This declaration also indicates that given the relative difference in expression levels, the disclosed nucleic acids and corresponding polypeptides have utility as cancer diagnostic tools. The Examiner has not offered any substantial reason or evidence to question the data in Example 18, or the first Grimaldi Declaration.

Second, Applicants have demonstrated that it is not necessary to know the cause or consequence of the differential expression of PRO1027 nucleic acids and polypeptides in certain tumors in order to use them as diagnostic tools for cancer.

Third, Applicants have shown that the second Grimaldi Declaration and Polakis Declaration, the accompanying references, as well as the excerpts and references cited above, demonstrate that it is well-established in the art that a change in mRNA levels generally correlates to a corresponding change in the encoded protein levels. The Examiner has not offered any substantial reason or evidence to question these declarations and supporting references. To the contrary, the Examiner has cited references that support the principle that mRNA levels generally correlate with protein levels.

Finally, the Examiner asserts that there is no asserted specific utility. Applicants have pointed out that the substantial utilities described above are specific to the claimed polypeptides because the PRO1027 gene and polypeptide are differentially expressed in melanoma cells compared to the corresponding normal skin cells. This is not a general utility that would apply to the broad class of polypeptides.

Given the totality of the evidence provided, Applicants submit that they have established a substantial, specific, and credible utility for the claimed polypeptides as diagnostic tools. According to the PTO Utility Examination Guidelines (2001), irrefutable proof of a claimed utility is not required. Rather, a specific, substantial, and credible utility requires only a "reasonable" confirmation of a real world context of use. Applicants remind the Examiner that:

> A small degree of utility is sufficient . . . The claimed invention must only be capable of performing some beneficial function . . . An invention does not lack utility merely because the particular embodiment disclosed in the patent lacks perfection or performs crudely . . . A commercially successful product is not required . . . Nor is it essential that the invention accomplish all its intended functions . . . or operate under all conditions . . . partial success being sufficient to demonstrate patentable utility . . . In short, **the defense of non-utility cannot be sustained without proof of total incapacity**. If an invention is only partially successful in achieving a useful result, a rejection of the claimed invention as a

whole based on a lack of utility is not appropriate. M.P.E.P. at 2107.01 (underline emphasis in original, bold emphasis added, citations omitted).

Applicants submit that they have established that it is more likely than not that one of skill in the art would reasonably accept the utility for the claimed polypeptides relating to PRO1027 set forth in the specification. In view of the above, Applicants respectfully request that the PTO reconsider and withdraw the utility rejection under 35 U.S.C. §101.

Rejection under 35 U.S.C. § 112, first paragraph – Enablement

The Examiner maintains the rejection of Claims 1-8 and 11-13 as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to use the invention. The Examiner maintains that because the claimed invention is not supported by either a specific, substantial and credible asserted utility, or a well-established utility, the claims are not enabled.

Applicants submit that in the discussion of the rejection under 35 U.S.C. §101 above, Applicants have established a substantial, specific and credible utility for the claimed polypeptides. Applicants therefore request that the Examiner withdraw the enablement rejection under 35 U.S.C. §112, first paragraph.

The Examiner has also rejected Claims 1-6 and 11-13 on the grounds that the specification allegedly lacks compete deposit information for the deposit of the full-length cDNA encoding the claimed polypeptide. The Examiner requests that Applicants provide a declaration or affidavit providing assurances that all restrictions upon public access to the ATCC™ accession number 203245 as specifically claimed will be irrevocably removed upon the grant of a patent. Although paragraph [446] of the specification states that permanent and unrestricted availability of the deposits will be provided upon issuance of the pertinent U.S. Patent or laying open to the public of any U.S. or foreign patent application, whichever comes first, Applicants provide the requested Declaration herewith. Applicants request that the Examiner withdraw the enablement rejection accordingly.

Rejection under 35 U.S.C. § 112, first paragraph – Written Description

The Examiner has maintained the rejection of Claims 1-5 and 12-13 under 35 U.S.C. § 112, first paragraph, as allegedly containing subject matter which was not described in the

specification in such a way as to reasonably convey to one skilled in the relevant art that the Applicants, at the time the application was filed, had possession of the claimed invention. The Examiner asserts that the addition of the functional limitation "wherein said isolated nucleic acid encodes a polypeptide that is more highly expressed in melanoma relative to normal skin cells" does not describe the function of the protein and does not impart specific structural information. The Examiner also maintains that the specification does not teach variants of either the nucleic acid or polypeptide. According to the Examiner, the skilled artisan would not be able to readily envision the claimed genus. Further, the Examiner asserts that while the skilled artisan may envision changes to the polypeptide of SEQ ID NO: 56, one cannot envision what changes to the structure or what part of the structure is conserved.

The Legal Standard for Written Description

The well-established test for sufficiency of support under the written description requirement of 35 U.S.C. §112 , first paragraph is whether the disclosure "reasonably conveys to artisan that the inventor had possession at that time of the later claimed subject matter." *In re Kaslow*, 707 F.2d 1366, 1375, 2121 USPQ 1089, 1096 (Fed. Cir. 1983); *see also Vas-Cath, Inc. v. Mahurkar*, 935 F.2d at1563, 19 USPQ2d at 1116 (Fed. Cir. 1991). The adequacy of written description support is a factual issue and is to be determined on a case-by-case basis. *See e.g., Vas-Cath, Inc. v. Mahurkar*, 935 F.2d at 1563, 19 USPQ2d at 1116 (Fed. Cir. 1991). The factual determination in a written description analysis depends on the nature of the invention and the amount of knowledge imparted to those skilled in the art by the disclosure. *Union Oil v. Atlantic Richfield Co.*, 208 F.3d 989, 996 (Fed. Cir. 2000).

The Current Invention is Adequately Described

As noted above, whether the Applicants were in possession of the invention as of the effective filing date of an application is a factual determination, reached by the consideration of a number of factors, including the level of knowledge and skill in the art, and the teaching provided by the specification. The inventor is not required to describe every single detail of his/her invention. An Applicant's disclosure obligation varies according to the art to which the invention pertains. The present invention pertains to the field of recombinant DNA/protein technology. It is well-established that the level of skill in this field is very high since a representative person of skill is generally a Ph.D. scientist with several years of experience.

Accordingly, the teaching imparted in the specification must be evaluated through the eyes of a highly skilled artisan as of the date the invention was made.

As amended, the pending claims are related to isolated polypeptides having at least 95% or 99% amino acid sequence identity to several polypeptides related to SEQ ID NO: 56, and satisfy the limitation "wherein said isolated polypeptide is more highly expressed in melanoma compared to normal skin, or wherein said isolated polypeptide is encoded by a polynucleotide that is more highly expressed in melanoma compared to normal skin" or "wherein said isolated polypeptide or a fragment thereof can be used to generate an antibody which can be used to specifically detect the polypeptide of SEQ ID NO: 56 in skin tissue samples."

Applicants maintain that there is no substantial variation within the species which fall within the scope of the amended claims, which require at least 95% or 99% amino acid sequence identity to the disclosed sequences related to SEQ ID NO: 56. Applicants note that the pending Claims are analogous to the claims discussed in Example 14 of the written description training materials. In Example 14, the written description requirement was found to be satisfied for claims relating to polypeptides having 95% homology to a particular sequence and possessing a particular catalytic activity, even though the applicant had not made any variants. Similarly, the pending claims also have very high sequence homology to the disclosed sequences and must share the same expression pattern in certain tumors, or share an epitope sufficient to generate antibodies which specifically detect the polypeptide of SEQ ID NO: 56 in skin tissue samples.

In Example 14, the procedures for making variants were known in the art and the disclosure taught how to test for the claimed catalytic activity. Similarly, in the instant application, it is well known in the art how to make polypeptides with at least 95% amino acid sequence identity to the disclosed sequences. In addition, the specification discloses how to test to determine if the polypeptide or encoding nucleic acid is differentially expressed in melanoma, and how to make antibodies which specifically detect the polypeptide of SEQ ID NO: 56 in skin tissue samples. Like Example 14, the genus of polypeptides that have at least 95% or 99% amino acid sequence identity to the disclosed sequences will not have substantial variation.

Furthermore, while Applicants appreciate that actions taken by the PTO in other applications are not binding with respect to the examination of the present application, Applicants note that the PTO has issued many patents containing claims to variant nucleic acids or variant proteins where the applicants did not actually make such nucleic acids or proteins.

Representative patents include U.S. Patent No: 6,737,522, U.S. Patent No. 6,395,306, U.S. Patent No. 6,025,156, U.S. Patent No. 6,645,499, U.S. Patent No. 6,498,235, and U.S. Patent No. 6,730,502 which are attached hereto as Exhibits 10-15.

In conclusion, Applicants submit that they have satisfied the written description requirement for the pending claims based on the actual reduction to practice of SEQ ID NO: 56, by specifying a high level of amino acid sequence identity, by describing how to test for differential expression of the polypeptide and encoding nucleic acid, and by describing how to make antibodies to the disclosed sequence, all of which result in a lack of substantial variability in the species falling within the scope of the instant claims. Applicants submit that this disclosure would allow one of skill in the art to "recognize that the applicant was in possession of the necessary common attributes or features of the elements possessed by the members of the genus." Hence, Applicants respectfully request that the PTO reconsider and withdraw the written description rejection under 35 U.S.C. §112.

Rejection under 35 U.S.C. § 102(b)

The Examiner has maintained the rejection of Claims 1-7, 11 and 12 under 35 U.S.C. § 102(b) as allegedly being anticipated by Rhodes et al. (1999). The Examiner has maintained the position that Applicants are not entitled to the claimed priority dates of June 24, 1998 or August 24, 2000, because the priority documents allegedly lack written description, enablement and utility for the reasons discussed above. The Examiner alleges that the Stemple Doctrine cannot be used to overcome the rejection set forth under 35 U.S.C. § 102(b).

Applicants respectfully disagree. As previously noted, in *In re Moore*, which followed the rule set forth in *In re Stemple*, the CCPA expressly stated that "the determination of a practical utility when one is not obvious need not have been accomplished prior to the date of a reference unless the reference also teaches how to use the compound it describes." *In re Moore*, 170 USPQ 260, 267 (CCPA 1971). The Rhodes et al. May, 1999 sequence database submission that forms the basis of the Examiner's rejection disclosed amino acids 1-77 of SEQ ID NO: 56. The Examiner has maintained that the disclosure of SEQ ID NO: 56 does not comply with the utility or enablement requirements. In other words, under the Examiner's reasoning, Rhodes et al. does "not teach how to use the compound it describes." Accordingly, under *Moore*, since Rhodes et al. does not teach how to use SEQ ID NO: 56, Applicants' disclosure in 60/090,444, filed **June 24, 1998**, cannot be dismissed as a priority document due to its alleged lack of utility.

Applicants have demonstrated, by means of the disclosure in 60/090,444 that they were in possession of so much of the claimed invention, i.e., SEQ ID NO: 56, as disclosed in the Rhodes reference dated May 1, 1999. Thus, Applicants submit that the cited reference is not available as prior art under 35 U.S.C. § 102(b), and request that the rejection be withdrawn.

CONCLUSION

In view of the above, Applicants respectfully maintain that claims are patentable and request that they be passed to issue. Applicants invite the Examiner to call the undersigned if any remaining issues may be resolved by telephone.

Please charge any additional fees, including any fees for additional extension of time, or credit overpayment to Deposit Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

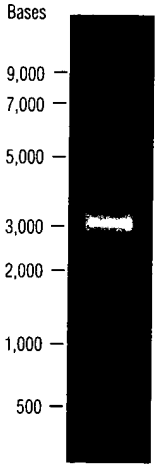
Dated: March 16, 2005

By: _____
AnneMarie Kaiser
Registration No. 37,649
Attorney of Record
Customer No. 30,313
(619) 235-8550

1327167_1
030805

RNA Ladder

#N0362S 25 µg $55



1 µg of RNA Ladder was heated at 60°C for 3 minutes in 1X RNA Ladder Sample Buffer and visualized by ethidium bromide staining (1.0% TBE agarose gel).

Description: The RNA Ladder is a set of 7 RNA molecules produced by *in vitro* transcription of a mixture of 7 linear DNA templates. The ladder sizes are: 9000, 7000, 5000, 3000, 2000, 1000 and 500 bases. The 3000 base fragment is at double intensity to serve as a reference band. This ladder is suitable for use as an RNA size standard on denaturing or native agarose gels.

Reagents Supplied with Ladders:
2X RNA Ladder Sample Buffer (for use with native agarose gels)

2X RNA Ladder Sample Buffer:
2X TBE (pH 8.3), 13% ficoll (w/v), 0.01% bromophenol blue and 7 M urea.

Concentration: 500 µg/ml.

Storage Conditions: 20 mM KOAc (pH 4.5). Store at –70°C. For short term storage (< 1 week), ladder can be stored at –20°C.

Notes on Use:
To avoid ribonuclease contamination: wear gloves, use RNase-free water for gels and buffers, wash equipment with detergent and rinse thoroughly with RNase-free water.

It is best to use freshly poured gels that are as thin as possible (i.e, 2–10 mm). Excessively long run times or high voltage can cause degradation of the bands on the gel. We recommend 4–8 volts/cm and running the bromophenol blue approximately 5 cm into the gel for good resolution.

Adding ethidium bromide to agarose gels and running buffer at a final concentration of 0.5 µg/ml will effectively stain the bands during electrophoresis.

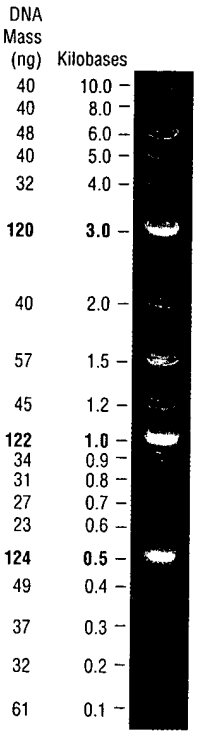
Denaturing vs. Native Agarose Gels:
It is common practice to electrophorese RNA on a fully denaturing agarose gel, such as one containing formaldehyde (1). However, in many cases it is possible to run RNA on a native agarose gel and obtain suitable results. In fact, it has been demonstrated that treatment of RNA samples in a denaturing buffer maintains the RNA molecules in a denatured state, during electrophoresis, for at least 3 hours (2,3). The use of native agarose gels eliminates problems associated with toxic chemicals and the difficulties encountered when staining and blotting formaldehyde gels.

References:
(1) Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, (2nd ed.), (pp. 7.43–7.45). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
(2) Liu, Y-C., Chou, Y-C. (1990) *Biotechniques* 9, 558.
(3) Sandra Cook and Christina Marchetti, unpublished observations.

2-Log DNA Ladder (0.1–10.0 kb)

New

#N3200S 100 µg $55
#N3200L 500 µg $220



1.0 µg of 2-Log DNA Ladder visualized by ethidium bromide staining on a 1.0% TBE agarose gel.

Description: A number of proprietary plasmids are digested to completion with appropriate restriction enzymes to yield 19 bands suitable for use as molecular weight standards for agarose gel electrophoresis. This digested DNA includes fragments ranging from 100 bp to 10 kb. The 0.5, 1.0 and 3.0 kb bands have increased intensity to serve as reference points.

Preparation: Double-stranded DNA is digested to completion with the appropriate restriction enzymes, phenol extracted and equilibrated to 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA.

Concentration: 1,000 µg/ml.

Storage Conditions: 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA. For long term storage, store at –20°C. 2-Log DNA Ladder is stable for at least 3 months at 4°C.

Note: All fragments have 4-base, 5′ overhangs that can be end labeled using T4 Polynucleotide Kinase (NEB #M0201) or filled-in using DNA Polymerase I, Klenow Fragment (NEB #M0210) (1). Use α-[^{32}P] dATP or α-[^{32}P] dTTP for the fill-in reaction.

Usage Recommendation: We recommend loading 1 µg of the 2-Log DNA Ladder diluted in sample buffer. This ladder was not designed for precise quantification of DNA mass but can be used for approximating the mass of DNA in comparably intense samples of similar size.

The approximate mass of DNA in each of the bands in our 2-Log DNA Ladder is as follows (assuming a 1 µg loading):

Fragment	Base Pairs	DNA Mass
1	10,002	40 ng
2	8,001	40 ng
3	6,001	48 ng
4	5,001	40 ng
5	4,001	32 ng
6	**3,001**	**120 ng**
7	2,017	40 ng
8	1,517	57 ng
9	1,200	45 ng
10	**1,000**	**122 ng**
11	900	34 ng
12	800	31 ng
13	700	27 ng
14	600	23 ng
15a	**517** >	**124 ng**
15b	**500**	
16	400	49 ng
17	300	37 ng
18	200	32 ng
19	100	61 ng

Reference:
(1) Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, (2nd ed.), (pp. 10.51–10.67). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

EXHIBIT 2

MOLECULAR BIOLOGY OF THE CELL

THIRD EDITION

Text Editor: Miranda Robertson
Managing Editor: Ruth Adams
Illustrator: Nigel Orme
Molecular Model Drawings: Kate Hesketh-Moore
Director of Electronic Publishing: John M-Roblin
Computer Specialist: Chuck Bartelt
Disk Preparation: Carol Winter
Copy Editor: Shirley M. Cobert
Production Editor: Douglas Goertzen
Production Coordinator: Perry Bessas
Indexer: Maija Hinkle

Bruce Alberts received his Ph.D. from Harvard University and is currently President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. *Dennis Bray* received his Ph.D. from the Massachusetts Institute of Technology and is currently a Medical Research Council Fellow in the Department of Zoology, University of Cambridge. *Julian Lewis* received his D.Phil. from the University of Oxford and is currently a Senior Scientist in the Imperial Cancer Research Fund Developmental Biology Unit, University of Oxford. *Martin Raff* received his M.D. from McGill University and is currently a Professor in the MRC Laboratory for Molecular Cell Biology and the Biology Department, University College London. *Keith Roberts* received his Ph.D. from the University of Cambridge and is currently Head of the Department of Cell Biology, the John Innes Institute, Norwich. *James D. Watson* received his Ph.D. from Indiana University and is currently Director of the Cold Spring Harbor Laboratory. He is the author of *Molecular Biology of the Gene* and, with Francis Crick and Maurice Wilkins, won the Nobel Prize in Medicine and Physiology in 1962.

Library of Congress Cataloging-in-Publication Data
Molecular biology of the cell / Bruce Alberts . . . [et al.].—3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-8153-1619-4 (hard cover).—ISBN 0-8153-1620-8 (pbk.)
1. Cytology. 2. Molecular biology. I. Alberts, Bruce.
[DNLM: 1. Cells. 2. Molecular Biology. QH 581.2 M718 1994]
QH581.2.M64 1994
574.87—dc20
DNLM/DLC
for Library of Congress 93-45907
CIP

Published by Garland Publishing, Inc.
717 Fifth Avenue, New York, NY 10022

Printed in the United States of America

15 14 13 12 10 9 8 7

Front cover: The photograph shows a rat nerve cell in culture. It is labeled (*yellow*) with a fluorescent antibody that stains its cell body and dendritic processes. Nerve terminals (*green*) from other neurons (not visible), which have made synapses on the cell, are labeled with a different antibody. (Courtesy of Olaf Mundigl and Pietro de Camilli.)

Dedication page: Gavin Borden, late president of Garland Publishing, weathered in during his mid-1980s climb near Mount McKinley with MBoC author Bruce Alberts and famous mountaineer guide Mugs Stump (1940–1992).

Back cover: The authors, in alphabetical order, crossing Abbey Road in London on their way to lunch. Much of this third edition was written in a house just around the corner. (Photograph by Richard Olivier.)

extracts. If these minor cell proteins differ among cells to the same extent as the more abundant proteins, as is commonly assumed, only a small number of protein differences (perhaps several hundred) suffice to create very large differences in cell morphology and behavior.

A Cell Can Change the Expression of Its Genes in Response to External Signals [3]

Most of the specialized cells in a multicellular organism are capable of altering their patterns of gene expression in response to extracellular cues. If a liver cell is exposed to a glucocorticoid hormone, for example, the production of several specific proteins is dramatically increased. Glucocorticoids are released during periods of starvation or intense exercise and signal the liver to increase the production of glucose from amino acids and other small molecules; the set of proteins whose production is induced includes enzymes such as tyrosine aminotransferase, which helps to convert tyrosine to glucose. When the hormone is no longer present, the production of these proteins drops to its normal level.

Other cell types respond to glucocorticoids in different ways. In fat cells, for example, the production of tyrosine aminotransferase is reduced, while some other cell types do not respond to glucocorticoids at all. These examples illustrate a general feature of cell specialization—different cell types often respond in different ways to the same extracellular signal. Underlying this specialization are features that do not change, which give each cell type its permanently distinctive character. These features reflect the persistent expression of different sets of genes.

Gene Expression Can Be Regulated at Many of the Steps in the Pathway from DNA to RNA to Protein [4]

If differences between the various cell types of an organism depend on the particular genes that the cells express, at what level is the control of gene expression exercised? There are many steps in the pathway leading from DNA to protein, and all of them can in principle be regulated. Thus a cell can control the proteins it makes by (1) controlling when and how often a given gene is transcribed (**transcriptional control**), (2) controlling how the primary RNA transcript is spliced or otherwise processed (**RNA processing control**), (3) selecting which completed mRNAs in the cell nucleus are exported to the cytoplasm (**RNA transport control**), (4) selecting which mRNAs in the cytoplasm are translated by ribosomes (**translational control**), (5) selectively destabilizing certain mRNA molecules in the cytoplasm (**mRNA degradation control**), or (6) selectively activating, inactivating, or compartmentalizing specific protein molecules after they have been made (**protein activity control**) (Figure 9–2).

For most genes transcriptional controls are paramount. This makes sense because, of all the possible control points illustrated in Figure 9–2, only transcriptional control ensures that no superfluous intermediates are synthesized. In the

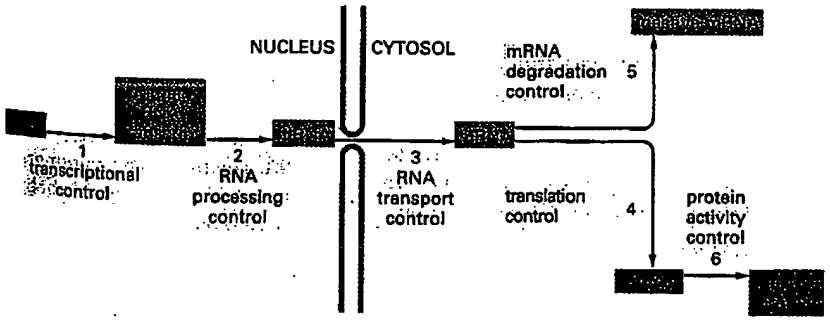


Figure 9–2 Six steps at which eucaryote gene expression can be controlled. Only controls that operate at steps 1 through 5 are discussed in this chapter. The regulation of protein activity (step 6) is discussed in Chapter 5; this includes reversible activation or inactivation by protein phosphorylation as well as irreversible inactivation by proteolytic degradation.

following sections we discuss the DNA and protein components that regulate the initiation of gene transcription. We return at the end of the chapter to the other ways of regulating gene expression.

Summary

The genome of a cell contains in its DNA sequence the information to make many thousands of different protein and RNA molecules. A cell typically expresses only a fraction of its genes, and the different types of cells in multicellular organisms arise because different sets of genes are expressed. Moreover, cells can change the pattern of genes they express in response to changes in their environment, such as signals from other cells. Although all of the steps involved in expressing a gene can in principle be regulated, for most genes the initiation of RNA transcription is the most important point of control.

DNA-binding Motifs in Gene Regulatory Proteins [5]

How does a cell determine which of its thousands of genes to transcribe? As discussed in Chapter 8, the transcription of each gene is controlled by a regulatory region of DNA near the site where transcription begins. Some regulatory regions are simple and act as switches that are thrown by a single signal. Other regulatory regions are complex and act as tiny microprocessors, responding to a variety of signals that they interpret and integrate to switch the neighboring gene on or off. Whether complex or simple, these switching devices consist of two fundamental types of components: (1) short stretches of DNA of defined sequence and (2) *gene regulatory proteins* that recognize and bind to them.

We begin our discussion of gene regulatory proteins by describing how these proteins were discovered.

Gene Regulatory Proteins Were Discovered Using Bacterial Genetics [6]

Genetic analyses in bacteria carried out in the 1950s provided the first evidence of the existence of **gene regulatory proteins** that turn specific sets of genes on or off. One of these regulators, the *lambda repressor*, is encoded by a bacterial virus, *bacteriophage lambda*. The repressor shuts off the viral genes that code for the protein components of new virus particles and thereby enables the viral genome to remain a silent passenger in the bacterial chromosome, multiplying with the bacterium when conditions are favorable for bacterial growth (see Figure 6–80). The lambda repressor was among the first gene regulatory proteins to be characterized, and it remains one of the best understood, as we discuss later. Other bacterial regulators respond to nutritional conditions by shutting off genes encoding specific sets of metabolic enzymes when they are not needed. The *lac repressor*, for example, the first of these bacterial proteins to be recognized, turns off the production of the proteins responsible for lactose metabolism when this sugar is absent from the medium.

The first step toward understanding gene regulation was the isolation of mutant strains of bacteria and bacteriophage lambda that were unable to shut off specific sets of genes. It was proposed at the time, and later proved, that most of these mutants were deficient in proteins acting as specific repressors for these sets of genes. Because these proteins, like most gene regulatory proteins, are present in small quantities, it was difficult and time-consuming to isolate them. They were eventually purified by fractionating cell extracts on a series of standard chromatography columns (see pp. 166–169). Once isolated, the proteins were shown to bind to specific DNA sequences close to the genes that they

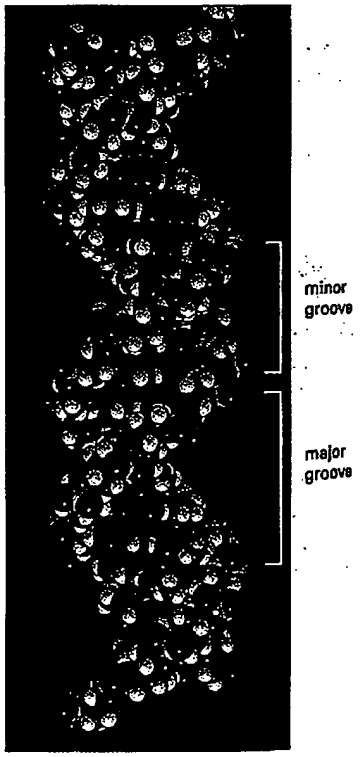


Figure 9–3 Double-helical structure of DNA. The major and minor grooves on the outside of the double helix are indicated. The atoms are colored as follows: carbon, *dark blue;* nitrogen, *light blue;* hydrogen, *white;* oxygen, *red;* phosphorus, *yellow.*

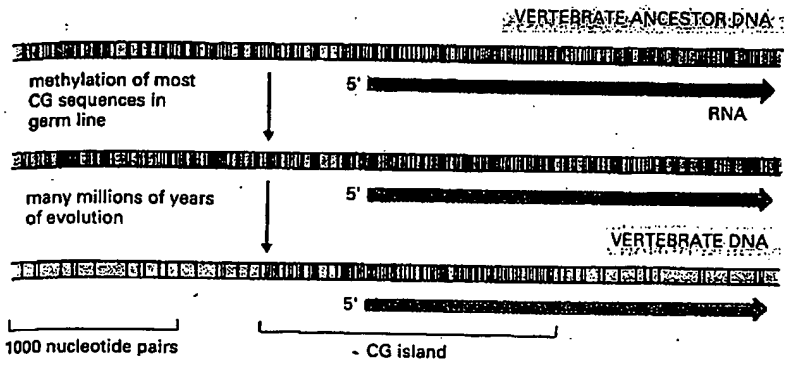


Figure 9–71 A mechanism to explain both the marked deficiency of CG sequences and the presence of CG islands in vertebrate genomes. A *black line* marks the location of an unmethylated CG dinucleotide in the DNA sequence, while a *red line* marks the location of a methylated CG dinucleotide.

Summary

The many types of cells in animals and plants are created largely through mechanisms that cause different genes to be transcribed in different cells. Since many specialized animal cells can maintain their unique character when grown in culture, the gene regulatory mechanisms involved in creating them must be stable once established and heritable when the cell divides, endowing the cell with a memory of its developmental history. Procaryotes and yeasts provide unusually accessible model systems in which to study gene regulatory mechanisms, some of which may be relevant to the creation of specialized cell types in higher eucaryotes. One such mechanism involves a competitive interaction between two (or more) gene regulatory proteins, each of which inhibits the synthesis of the other; this can create a flip-flop switch that switches a cell between two alternative patterns of gene expression. Direct or indirect positive feedback loops, which enable gene regulatory proteins to perpetuate their own synthesis, provide a general mechanism for cell memory.

In eucaryotes gene transcription is generally controlled by combinations of gene regulatory proteins. It is thought that each type of cell in a higher eucaryotic organism contains a specific combination of gene regulatory proteins that ensures the expression of only those genes appropriate to that type of cell. A given gene regulatory protein may be expressed in a variety of circumstances and typically is involved in the regulation of many genes.

In addition to diffusible gene regulatory proteins, inherited states of chromatin condensation are also utilized by eucaryotic cells to regulate gene expression. In vertebrates DNA methylation also plays a part, mainly as a device to reinforce decisions about gene expression that are made initially by other mechanisms.

Posttranscriptional Controls

Although controls on the initiation of gene transcription are the predominant form of regulation for most genes, other controls can act later in the pathway from RNA to protein to modulate the amount of gene product that is made. Although these **posttranscriptional controls**, which operate after RNA polymerase has bound to the gene's promoter and begun RNA synthesis, are less common than *transcriptional control*, for many genes they are crucial. It seems that every step in gene expression that could be controlled in principle is likely to be regulated under some circumstances for some genes.

We consider the varieties of posttranscriptional regulation in temporal order, according to the sequence of events that might be experienced by an RNA molecule after its transcription has begun (Figure 9–72).

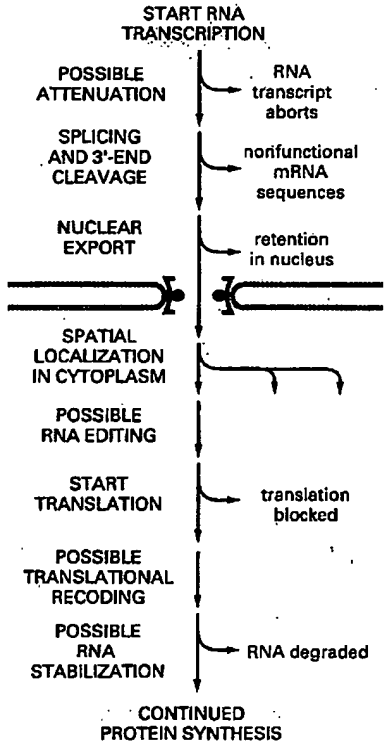


Figure 9–72 Possible posttranscriptional controls on gene expression. Only a few of these controls are likely to be used for any one gene.

EXHIBIT 3

MOLECULAR BIOLOGY OF THE CELL

fourth edition

Bruce Alberts

Alexander Johnson

Julian Lewis

Martin Raff

Keith Roberts

Peter Walter

GS Garland Science
Taylor & Francis Group

Garland
Vice President: Denise Schanck
Managing Editor: Sarah Gibbs
Senior Editorial Assistant: Kirsten Jenner
Managing Production Editor: Emma Hunt
Proofreader and Layout: Emma Hunt
Production Assistant: Angela Bennett
Text Editors: Marjorie Singer Anderson and Betsy Dilernia
Copy Editor: Bruce Goatly
Word Processors: Fran Dependahl, Misty Landers and Carol Winter
Designer: Blink Studio, London
Illustrator: Nigel Orme
Indexer: Janine Ross and Sherry Granum
Manufacturing: Nigel Eyre and Marion Morrow

Bruce Alberts received his Ph.D. from Harvard University and is President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. **Alexander Johnson** received his Ph.D. from Harvard University and is a Professor of Microbiology and Immunology at the University of California, San Francisco. **Julian Lewis** received his D.Phil. from the University of Oxford and is a Principal Scientist at the Imperial Cancer Research Fund, London. **Martin Raff** received his M.D. from McGill University and is at the Medical Research Council Laboratory for Molecular Cell Biology and Cell Biology Unit and in the Biology Department at University College London. **Keith Roberts** received his Ph.D. from the University of Cambridge and is Associate Research Director at the John Innes Centre, Norwich. **Peter Walter** received his Ph.D. from The Rockefeller University in New York and is Professor and Chairman of the Department of Biochemistry and Biophysics at the University of California, San Francisco, and an Investigator of the Howard Hughes Medical Institute.

Library of Congress Cataloging-in-Publicaton Data
Molecular biology of the cell / Bruce Alberts ... [et al.].-- 4th ed.
p. cm
Includes bibliographical references and index.
ISBN 0-8153-3218-1 (hardbound) -- ISBN 0-8153-4072-9 (pbk.)
1. Cytology. 2. Molecular biology. I. Alberts, Bruce.
[DNLM: 1. Cells. 2. Molecular Biology.]
QH581.2 .M64 2002
571.6--dc21
2001054471 CIP

Published by Garland Science, a member of the Taylor & Francis Group, 29 West 35th Street, New York, NY 10001-2299

Printed in the United States of America

15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Cell Biology Interactive
Artistic and Scientific Direction: Peter Walter
Narrated by: Julie Theriot
Production, Design, and Development: Mike Morales

Front cover Human Genome: Reprinted by permission from *Nature*, International Human Genome Sequencing Consortium, 409:860–921, 2001 © Macmillan Magazines Ltd. Adapted from an image by Francis Collins, NHGRI; Jim Kent, UCSC; Ewan Birney, EBI; and Darryl Leja, NHGRI; showing a portion of Chromosome 1 from the initial sequencing of the human genome.

Back cover In 1967, the British artist Peter Blake created a design classic. Nearly 35 years later Nigel Orme (illustrator), Richard Denyer (photographer), and the authors have together produced an affectionate tribute to Mr Blake's image. With its gallery of icons and influences, its assembly created almost as much complexity, intrigue and mystery as the original. *Drosophila, Arabidopsis,* Dolly and the assembled company tempt you to dip inside where, as in the original, "a splendid time is guaranteed for all." (Gunter Blobel, courtesy of The Rockefeller University; Marie Curie, Keystone Press Agency Inc; Darwin bust, by permission of the President and Council of the Royal Society; Rosalind Franklin, courtesy of Cold Spring Harbor Laboratory Archives; Dorothy Hodgkin, © The Nobel Foundation, 1964; James Joyce, etching by Peter Blake; Robert Johnson, photo booth self-portrait early 1930s, © 1986 Delta Haze Corporation all rights reserved, used by permission; Albert L. Lehninger, (unidentified photographer) courtesy of The Alan Mason Chesney Medical Archives of The Johns Hopkins Medical Institutions; Linus Pauling, from Ava Helen and Linus Pauling Papers, Special Collections, Oregon State University; Nicholas Poussin, courtesy of ArtToday.com; Barbara McClintock, © David Micklos, 1983; Andrei Sakharov, courtesy of Elana Bonner; Frederick Sanger, © The Nobel Foundation, 1958.)

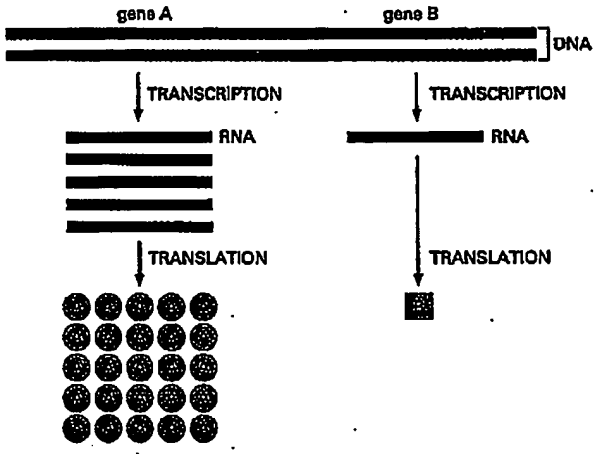


Figure 6–3 Genes can be expressed with different efficiencies. Gene A is transcribed and translated much more efficiently than gene B. This allows the amount of protein A in the cell to be much greater than that of protein B.

FROM DNA TO RNA

Transcription and translation are the means by which cells read out, or express, the genetic instructions in their genes. Because many identical RNA copies can be made from the same gene, and each RNA molecule can direct the synthesis of many identical protein molecules, cells can synthesize a large amount of protein rapidly when necessary. But each gene can also be transcribed and translated with a different efficiency, allowing the cell to make vast quantities of some proteins and tiny quantities of others (Figure 6–3). Moreover, as we see in the next chapter, a cell can change (or regulate) the expression of each of its genes according to the needs of the moment—most obviously by controlling the production of its RNA.

Portions of DNA Sequence Are Transcribed into RNA

The first step a cell takes in reading out a needed part of its genetic instructions is to copy a particular portion of its DNA nucleotide sequence—a gene—into an RNA nucleotide sequence. The information in RNA, although copied into another chemical form, is still written in essentially the same language as it is in DNA—the language of a nucleotide sequence. Hence the name **transcription.**

Like DNA, RNA is a linear polymer made of four different types of nucleotide subunits linked together by phosphodiester bonds (Figure 6–4). It differs from DNA chemically in two respects: (1) the nucleotides in RNA are *ribonucleotides*—that is, they contain the sugar ribose (hence the name *ribo*nucleic acid) rather than deoxyribose; (2) although, like DNA, RNA contains the bases adenine (A), guanine (G), and cytosine (C), it contains the base uracil (U) instead of the thymine (T) in DNA. Since U, like T, can base-pair by hydrogen-bonding with A (Figure 6–5), the complementary base-pairing properties described for DNA in Chapters 4 and 5 apply also to RNA (in RNA, G pairs with C, and A pairs with U). It is not uncommon, however, to find other types of base pairs in RNA: for example, G pairing with U occasionally.

Despite these small chemical differences, DNA and RNA differ quite dramatically in overall structure. Whereas DNA always occurs in cells as a double-stranded helix, RNA is single-stranded. RNA chains therefore fold up into a variety of shapes, just as a polypeptide chain folds up to form the final shape of a protein (Figure 6–6). As we see later in this chapter, the ability to fold into complex three-dimensional shapes allows some RNA molecules to have structural and catalytic functions.

Transcription Produces RNA Complementary to One Strand of DNA

All of the RNA in a cell is made by DNA transcription, a process that has certain similarities to the process of DNA replication discussed in Chapter 5.

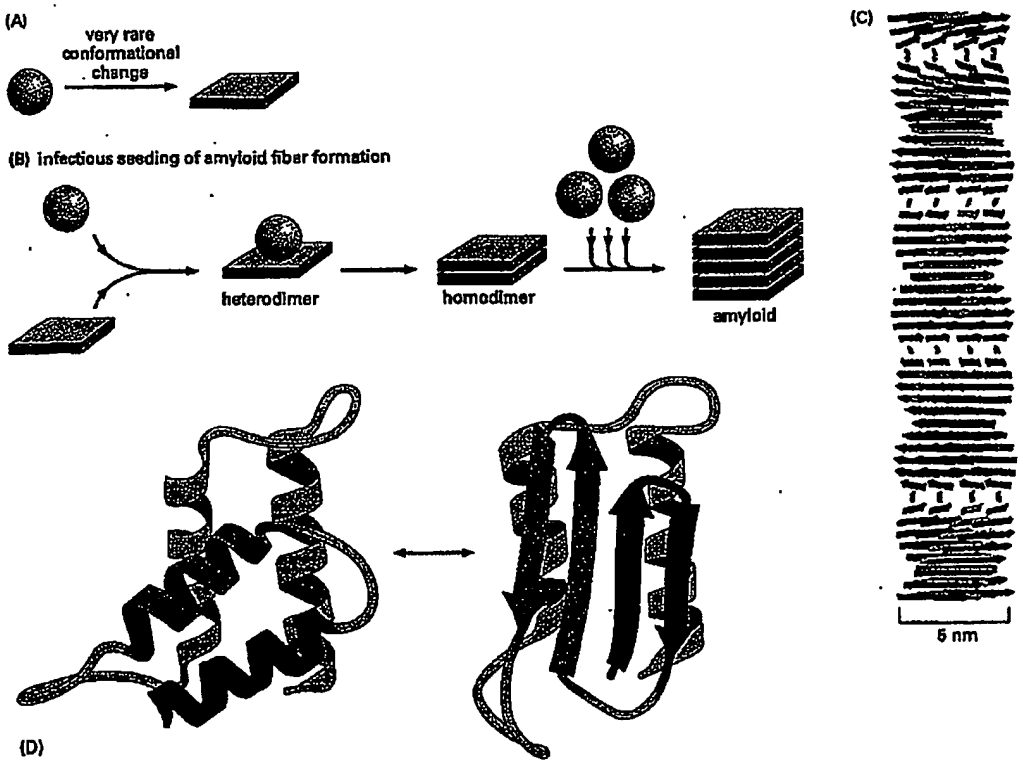


Figure 6–89 Protein aggregates that cause human disease. (A) Schematic illustration of the type of conformational change in a protein that produces material for a cross-beta filament. (B) Diagram illustrating the self-infectious nature of the protein aggregation that is central to prion diseases. PrP is highly unusual because the misfolded version of the protein, called PrP*, induces the normal PrP protein it contacts to change its conformation, as shown. Most of the human diseases caused by protein aggregation are caused by the overproduction of a variant protein that is especially prone to aggregation, but because this structure is not infectious in this way, it cannot spread from one animal to another. (C) Drawing of a cross-beta filament, a common type of protease-resistant protein aggregate found in a variety of human neurological diseases. Because the hydrogen-bond interactions in a β sheet form between polypeptide backbone atoms (see Figure 3–9), a number of different abnormally folded proteins can produce this structure. (D) One of several possible models for the conversion of PrP to PrP*, showing the likely change of two α-helices into four β-strands. Although the structure of the normal protein has been determined accurately, the structure of the infectious form is not yet known with certainty because the aggregation has prevented the use of standard structural techniques. (C, courtesy of Louise Serpell, adapted from M. Sunde et al., *J. Mol. Biol.* 273:729–739, 1997; D, adapted from S.B. Prusiner, *Trends Biochem. Sci.* 21:482–487, 1996.)

animals and humans. It can be dangerous to eat the tissues of animals that contain PrP*, as witnessed most recently by the spread of BSE (commonly referred to as the "mad cow disease") from cattle to humans in Great Britain.

Fortunately, in the absence of PrP*, PrP is extraordinarily difficult to convert to its abnormal form. Although very few proteins have the potential to misfold into an infectious conformation, a similar transformation has been discovered to be the cause of an otherwise mysterious "protein-only inheritance" observed in yeast cells.

There Are Many Steps From DNA to Protein

We have seen so far in this chapter that many different types of chemical reactions are required to produce a properly folded protein from the information contained in a gene (Figure 6–90). The final level of a properly folded protein in a cell therefore depends upon the efficiency with which each of the many steps is performed.

We discuss in Chapter 7 that cells have the ability to change the levels of their proteins according to their needs. In principle, any or all of the steps in Fig-

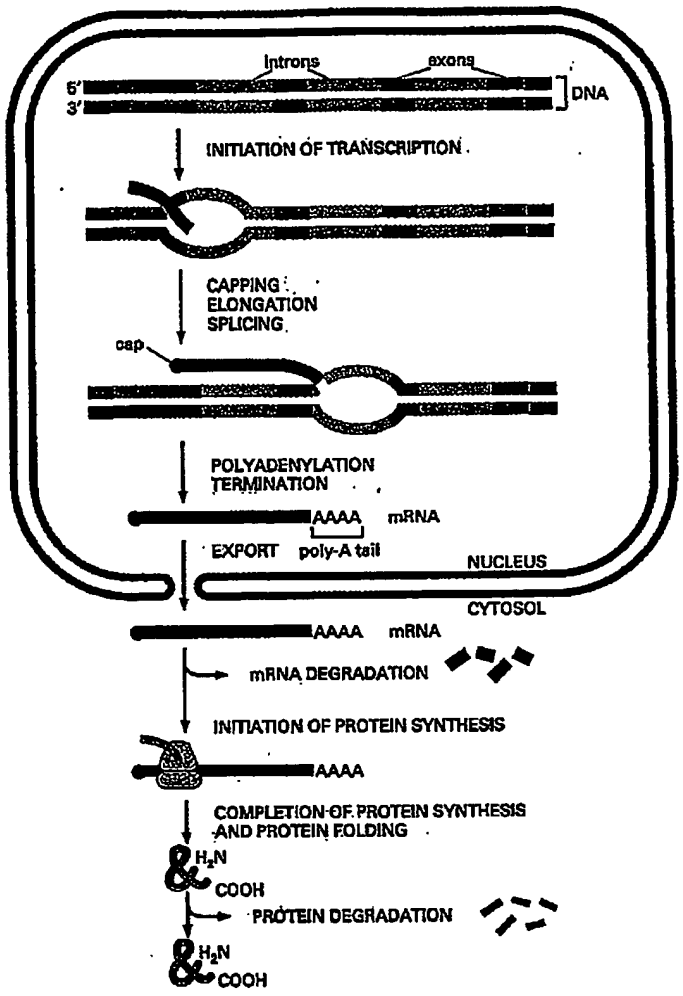


Figure 6–90 The production of a protein by a eucaryotic cell. The final level of each protein in a eucaryotic cell depends upon the efficiency of each step depicted.

ure 6–90) could be regulated by the cell for each individual protein. However, as we shall see in Chapter 7, the initiation of transcription is the most common point for a cell to regulate the expression of each of its genes. This makes sense, inasmuch as the most efficient way to keep a gene from being expressed is to block the very first step—the transcription of its DNA sequence into an RNA molecule.

Summary

The translation of the nucleotide sequence of an mRNA molecule into protein takes place in the cytoplasm on a large ribonucleoprotein assembly called a ribosome. The amino acids used for protein synthesis are first attached to a family of tRNA molecules, each of which recognizes, by complementary base-pair interactions, particular sets of three nucleotides in the mRNA (codons). The sequence of nucleotides in the mRNA is then read from one end to the other in sets of three according to the genetic code.

To initiate translation, a small ribosomal subunit binds to the mRNA molecule at a start codon (AUG) that is recognized by a unique initiator tRNA molecule. A large ribosomal subunit binds to complete the ribosome and begin the elongation phase of protein synthesis. During this phase, aminoacyl tRNAs—each bearing a specific amino acid bind sequentially to the appropriate codon in mRNA by forming complementary base pairs with the tRNA anticodon. Each amino acid is added to the C-terminal end of the growing polypeptide by means of a cycle of three sequential

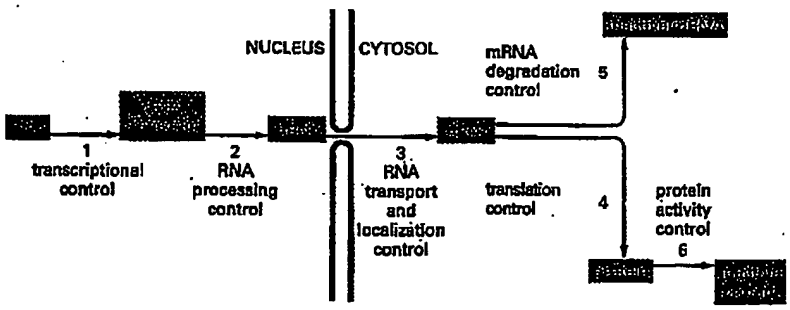


Figure 7–5 Six steps at which eucaryotic gene expression can be controlled. Controls that operate at steps 1 through 5 are discussed in this chapter. Step 6, the regulation of protein activity, includes reversible activation or inactivation by protein phosphorylation (discussed in Chapter 3) as well as irreversible inactivation by proteolytic degradation (discussed in Chapter 6).

Gene Expression Can Be Regulated at Many of the Steps in the Pathway from DNA to RNA to Protein

If differences among the various cell types of an organism depend on the particular genes that the cells express, at what level is the control of gene expression exercised? As we saw in the last chapter, there are many steps in the pathway leading from DNA to protein, and all of them can in principle be regulated. Thus a cell can control the proteins it makes by (1) controlling when and how often a given gene is transcribed (**transcriptional control**), (2) controlling how the RNA transcript is spliced or otherwise processed (**RNA processing control**), (3) selecting which completed mRNAs in the cell nucleus are exported to the cytosol and determining where in the cytosol they are localized (**RNA transport and localization control**), (4) selecting which mRNAs in the cytoplasm are translated by ribosomes (**translational control**), (5) selectively destabilizing certain mRNA molecules in the cytoplasm (**mRNA degradation control**), or (6) selectively activating, inactivating, degrading, or compartmentalizing specific protein molecules after they have been made (**protein activity control**) (Figure 7–5).

For most genes transcriptional controls are paramount. This makes sense because, of all the possible control points illustrated in Figure 7–5, only transcriptional control ensures that the cell will not synthesize superfluous intermediates. In the following sections we discuss the DNA and protein components that perform this function by regulating the initiation of gene transcription. We shall return at the end of the chapter to the additional ways of regulating gene expression.

Summary

The genome of a cell contains in its DNA sequence the information to make many thousands of different protein and RNA molecules. A cell typically expresses only a fraction of its genes, and the different types of cells in multicellular organisms arise because different sets of genes are expressed. Moreover, cells can change the pattern of genes they express in response to changes in their environment, such as signals from other cells. Although all of the steps involved in expressing a gene can in principle be regulated, for most genes the initiation of RNA transcription is the most important point of control.

DNA-BINDING MOTIFS IN GENE REGULATORY PROTEINS

How does a cell determine which of its thousands of genes to transcribe? As mentioned briefly in Chapters 4 and 6, the transcription of each gene is controlled by a regulatory region of DNA relatively near the site where transcription begins. Some regulatory regions are simple and act as switches that are thrown by a single signal. Many others are complex and act as tiny microprocessors, responding to a variety of signals that they interpret and integrate to switch the neighboring gene on or off. Whether complex or simple, these switching devices

occur in the germ line, the cell lineage that gives rise to sperm or eggs. Most of the DNA in vertebrate germ cells is inactive and highly methylated. Over long periods of evolutionary time, the methylated CG sequences in these inactive regions have presumably been lost through spontaneous deamination events that were not properly repaired. However promoters of genes that remain active in the germ cell lineages (including most housekeeping genes) are kept unmethylated, and therefore spontaneous deaminations of Cs that occur within them can be accurately repaired. Such regions are preserved in modern day vertebrate cells as CG islands. In addition, any mutation of a CG sequence in the genome that destroyed the function or regulation of a gene in the adult would be selected against, and some CG islands are simply the result of a higher than normal density of critical CG sequences.

The mammalian genome contains an estimated 20,000 CG islands. Most of the islands mark the 5′ ends of transcription units and thus, presumably, of genes. The presence of CG islands often provides a convenient way of identifying genes in the DNA sequences of vertebrate genomes.

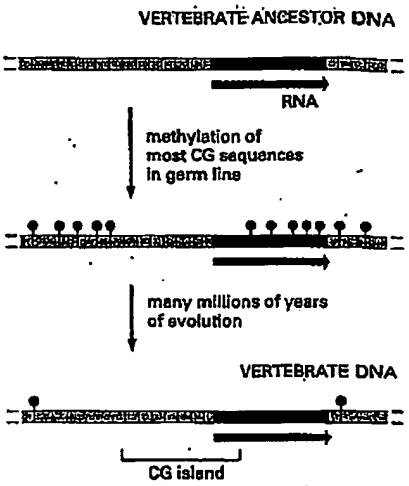


Figure 7–86 A mechanism to explain both the marked overall deficiency of CG sequences and their clustering into CG islands in vertebrate genomes. A *black line* marks the location of a CG dinucleotide in the DNA sequence, while a red "lollipop" indicates the presence of a methyl group on the CG dinucleotide. CG sequences that lie in regulatory sequences of genes that are transcribed in germ cells are unmethylated and therefore tend to be retained in evolution. Methylated CG sequences, on the other hand, tend to be lost through deamination of 5-methyl C to T, unless the CG sequence is critical for survival.

Summary

The many types of cells in animals and plants are created largely through mechanisms that cause different genes to be transcribed in different cells. Since many specialized animal cells can maintain their unique character through many cell division cycles and even when grown in culture, the gene regulatory mechanisms involved in creating them must be stable once established and heritable when the cell divides. These features endow the cell with a memory of its developmental history. Bacteria and yeasts provide unusually accessible model systems in which to study gene regulatory mechanisms. One such mechanism involves a competitive interaction between two gene regulatory proteins, each of which inhibits the synthesis of the other; this can create a flip-flop switch that switches a cell between two alternative patterns of gene expression. Direct or indirect positive feedback loops, which enable gene regulatory proteins to perpetuate their own synthesis, provide a general mechanism for cell memory. Negative feedback loops with programmed delays form the basis for cellular clocks.

In eucaryotes the transcription of a gene is generally controlled by combinations of gene regulatory proteins. It is thought that each type of cell in a higher eucaryotic organism contains a specific combination of gene regulatory proteins that ensures the expression of only those genes appropriate to that type of cell. A given gene regulatory protein may be active in a variety of circumstances and typically is involved in the regulation of many genes.

In addition to diffusible gene regulatory proteins, inherited states of chromatin condensation are also used by eucaryotic cells to regulate gene expression. An especially dramatic case is the inactivation of an entire X chromosome in female mammals. In vertebrates DNA methylation also functions in gene regulation, being used mainly as a device to reinforce decisions about gene expression that are made initially by other mechanisms. DNA methylation also underlies the phenomenon of genomic imprinting in mammals, in which the expression of a gene depends on whether it was inherited from the mother or the father.

POSTTRANSCRIPTIONAL CONTROLS

In principle, every step required for the process of gene expression could be controlled. Indeed, one can find examples of each type of regulation, although any one gene is likely to use only a few of them. Controls on the initiation of gene transcription are the predominant form of regulation for most genes. But other controls can act later in the pathway from DNA to protein to modulate the amount of gene product that is made. Although these **posttranscriptional controls**, which operate after RNA polymerase has bound to the gene's promoter and begun RNA synthesis, are less common than *transcriptional control*, for many genes they are crucial.

EXHIBIT 4

CHAPTER 29

Regulation of transcription

Genes VII (1997) CH 29, pp. 847-848.
Benjamin Lewin

The phenotypic differences that distinguish the various kinds of cells in a higher eukaryote are largely due to differences in the expression of genes that code for proteins, that is, those transcribed by RNA polymerase II. In principle, the expression of these genes might be regulated at any one of several stages. The concept of the "level of control" implies that gene expression is not necessarily an automatic process once it has begun. It could be regulated in a gene-specific way at any one of several sequential steps. We can distinguish (at least) five potential control points, forming the series:

Activation of gene structure
↓
Initiation of transcription
↓
Processing the transcript
↓
Transport to cytoplasm
↓
Translation of mRNA

The existence of the first step is implied by the discovery that genes may exist in either of two structural conditions. Relative to the state of most of the genome, genes are found in an "active" state in the cells in which they are expressed (see Chapter 27). The change of structure is distinct from the act of transcription, and indicates that the gene is "transcribable." This suggests that acquisition of the "active" structure must be the first step in gene expression.

Transcription of a gene in the active state is controlled at the stage of initiation, that is, by the interaction of RNA polymerase with its promoter. This is now becoming susceptible to analysis in the *in vitro* systems (see Chapter 28). For most genes, this is a major control point: probably it is the most common level of regulation.

There is at present no evidence for control at subsequent stages of transcription in eukaryotic cells, for example, via antitermination mechanisms.

The primary transcript is modified by capping at the 5' end, and usually also by polyadenylation at the 3' end. Introns must be spliced out from the transcripts of interrupted genes. The mature RNA must be exported from the nucleus to the cytoplasm. Regulation of gene expression by selection of sequences at the level of nuclear RNA might involve any or all of these stages, but the one for which we have most evidence concerns changes in splicing: some genes are expressed by means of alternative splicing patterns whose regulation controls the type of protein product (see Chapter 30).

Finally, the translation of an mRNA in the cytoplasm can be specifically controlled. There is little evidence for the employment of this mechanism in adult somatic cells, but it does occur in some embryonic situations, as described in Chapter 7. The mechanism is presumed to involve the blocking of initiation of translation of some mRNAs by specific protein factors.

But having acknowledged that control of gene expression can occur at multiple stages, and that production of RNA cannot inevitably be equated with production of protein, it is clear

that the overwhelming majority of regulatory events occur at the initiation of transcription. Regulation of tissue-specific gene transcription lies at the heart of eukaryotic differentiation; indeed, we see examples in Chapter 38 in which proteins that regulate embryonic development prove to be transcription factors. A regulatory transcription factor serves to provide common control of a large number of target genes, and we seek to answer two questions about this mode of regulation: what identifies the common target genes to the transcription factor; and how is the activity of the transcription factor itself regulated in response to intrinsic or extrinsic signals?

Response elements identify genes under common regulation

The principle that emerges from characterizing groups of genes under common control is that *they share a promoter element that is recognized by a regulatory transcription factor.* An element that causes a gene to respond to such a factor is called a response element; examples are the HSE (heat shock response element), GRE (glucocorticoid response element), SRE (serum response element).

The properties of some inducible transcription factors and the elements that they recognize are summarized in Table 29.1. Response elements have the same general characteristics as upstream elements of promoters or enhancers. They contain short consensus sequences, and copies of the response elements found in different genes are closely related, but not necessarily identical. The region bound by the factor extends for a short distance on either side of the consensus sequence. In promoters, the elements are not present at fixed distances from the startpoint, but are usually <200 bp upstream of it. The presence of a single element usually is sufficient to confer the regulatory response, but sometimes there are multiple copies.

Table 29.1 Inducible transcription factors bind to response elements that identify groups of promoters or enhancers subject to coordinate control.

Regulatory Agent	Module	Consensus	Factor
Heat shock	HSE	CNNGAANNTCCNNG	HSTF
Glucocorticoid	GRE	TGGTACAAATGTTCT	Receptor
Phorbol ester	TRE	TGACTCA	AP1
Serum	SRE	CCATATTAGG	SRF

Response elements may be located in promoters or in enhancers. Some types of elements are typically found in one rather than the other: usually an HSE is found in a promoter, while a GRE is found in an enhancer. We assume that all response elements function by the same general principle. *A gene is regulated by a sequence at the promoter or enhancer that is recognized by a specific protein. The protein functions as a transcription factor needed for RNA polymerase to initiate. Active protein is available only under conditions when the gene is to be expressed; its absence means that the promoter is not activated by this particular circuit.*

An example of a situation in which many genes are controlled by a single factor is provided by the heat shock response. This is common to a wide range of prokaryotes and eukaryotes and involves multiple controls of gene expression: an increase in temperature turns off transcription of some genes, turns on transcription of the heat shock genes, and causes changes in the translation of mRNAs. The control of the heat shock genes illustrates the differences between prokaryotic and eukaryotic modes of control. In bacteria, a new sigma factor is synthesized that directs RNA polymerase holoenzyme to recognize an alter-

EXHIBIT 5

World Journal of Surgical Oncology

BioMed Central

Research

Open Access

Prostate stem cell antigen (PSCA) expression in human prostate cancer tissues and its potential role in prostate carcinogenesis and progression of prostate cancer

Zhao Zhigang*[1] and Shen Wenlv[2]

Address: [1]Department of Urology, Shantou University Medical College, Shantou, Guangdong, China and [2]Department of Urology, No 2. Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong, China

Email: Zhao Zhigang* - zgzhaodr@yahoo.com.cn; Shen Wenlv - wlshen99@hotmail.com

* Corresponding author

Published: 10 May 2004

World Journal of Surgical Oncology 2004, 2:13

Received: 30 March 2004
Accepted: 10 May 2004

This article is available from: http://www.wjso.com/content/2/1/13

Abstract

Background: Prostate stem cell antigen (PSCA) is a recently defined homologue of the Thy-1/Ly-6 family of glycosylphosphatidylinositol (GPI)-anchored cell surface antigens. The purpose of the present study was to examine the expression status of PSCA protein and mRNA in clinical specimens of human prostate cancer (Pca) and to validate it as a potential molecular target for diagnosis and treatment of Pca.

Materials and Methods: Immunohistochemical (IHC) and *in situ* hybridization (ISH) analyses of PSCA expression were simultaneously performed on paraffin-embedded sections from 20 benign prostatic hyperplasia (BPH), 20 prostatic intraepithelial neoplasm (PIN) and 48 prostate cancer (Pca) tissues, including 9 androgen-independent prostate cancers. The level of PSCA expression was semiquantitatively scored by assessing both the percentage and intensity of PSCA-positive staining cells in the specimens. Then compared PSCA expression between BPH, PIN and Pca tissues and analysed the correlations of PSCA expression level with pathological grade, clinical stage and progression to androgen-independence in Pca.

Results: In BPH and low grade PIN, PSCA protein and mRNA staining were weak or negative and less intense and uniform than that seen in HGPIN and Pca. There were moderate to strong PSCA protein and mRNA expression in 8 of 11 (72.7%) HGPIN and in 40 of 48 (83.4%) Pca specimens examined by IHC and ISH analyses, with statistical significance compared with BPH (20%) and low grade PIN (22.2%) samples ($p < 0.05$, respectively). The expression level of PSCA increased with high Gleason grade, advanced stage and progression to androgen-independence ($p < 0.05$, respectively). In addition, IHC and ISH staining showed a high degree of correlation between PSCA protein and mRNA overexpression.

Conclusions: Our data demonstrate that PSCA as a new cell surface marker is overexpressed by a majority of human Pca. PSCA expression correlates positively with adverse tumor characteristics, such as increasing pathological grade (poor cell differentiation), worsening clinical stage and androgen-independence, and speculatively with prostate carcinogenesis. PSCA protein overexpression results from upregulated transcription of PSCA mRNA. PSCA may have prognostic utility and may be a promising molecular target for diagnosis and treatment of Pca.

Introduction

Prostate cancer (Pca) is the second leading cause of cancer-related death in American men and is becoming a common cancer increasing in China. Despite recently great progress in the diagnosis and management of localized disease, there continues to be a need for new diagnostic markers that can accurately discriminate between indolent and aggressive variants of Pca. There also continues to be a need for the identification and characterization of potential new therapeutic targets on Pca cells. Current diagnostic and therapeutic modalities for recurrent and metastatic Pca have been limited by a lack of specific target antigens of Pca.

Although a number of prostate-specific genes have been identified (i.e. prostate specific antigen, prostatic acid phosphatase, glandular kallikrein 2), the majority of these are secreted proteins not ideally suited for many immunological strategies. So, the identification of new cell surface antigens is critical to the development of new diagnostic and therapeutic approaches to the management of Pca.

Reiter RE et al [1] reported the identification of prostate stem cell antigen (PSCA), a cell surface antigen that is predominantly prostate specific. The PSCA gene encodes a 123 amino acid glycoprotein, with 30% homology to stem cell antigen 2 (Sca 2). Like Sca-2, PSCA also belongs to a member of the Thy-1/Ly-6 family and is anchored by a glycosylphosphatidylinositol (GPI) linkage. mRNA *in situ* hybridization (ISH) localized PSCA expression in normal prostate to the basal cell epithelium, the putative stem cell compartment of prostatic epithelium, suggesting that PSCA may be a marker of prostate stem/progenitor cells.

In order to examine the status of PSCA protein and mRNA expression in human Pca and validate it as a potential diagnostic and therapeutic target for Pca, we used immunohistochemistry (IHC) and *in situ* hybridization (ISH) simultaneously, and conducted PSCA protein and mRNA expression analyses in paraffin-embedded tissue specimens of benign prostatic hyperplasia (BPH, n = 20), prostate intraepithelial neoplasm (PIN, n = 20) and prostate cancer (Pca, n = 48). Furthermore, we evaluated the possible correlation of PSCA expression level with Pca tumorigenesis, grade, stage and progression to androgen-independence.

Materials and methods

Tissue samples

All of the clinical tissue specimens studied herein were obtained from 80 patients of 57–84 years old by prostatectomy, transurethral resection of prostate (TURP) or biopsies. The patients were classified as 20 cases of BPH, 20 cases of PIN, 40 cases of primary Pca, including 9 patients with recurrent Pca and a history of androgen ablation therapy (orchiectomy and/or hormonal therapy), who were referred to as androgen-independent prostate cancers. Eight specimens were harvested from these androgen-independent Pca patients prior to androgen ablation treatment. Each tissue sample was cut into two parts, one was fixed in 10% formalin for IHC and the other treated with 4% paraformaldehyde/0.1 M PBS PH 7.4 in 0.1% DEPC for 1 h for ISH analysis, and then embedded in paraffin. All paraffin blocks examined were then cut into 5 μm sections and mounted on the glass slides specific for IHC and ISH respectively in the usual fashion. H&E-stained section of each Pca was evaluated and assigned a Gleason score by the experienced urological pathologist at our institution based on the criteria of Gleason score [2]. The Gleason sums are summarized in Table 1. Clinical staging was performed according to Jewett-whitmore-prout staging system, as shown in Table 2. In the category of PIN, we graded the specimens into two groups, i.e. low grade PIN (grade I – II) and high grade PIN (HGPIN, grade III) on the basis of literatures [3,4].

Immunohistochemical (IHC) analysis

Briefly, tissue sections were deparaffinized, dehydrated, and subjected to microwaving in 10 mmol/L citrate buffer, PH 6.0 (Boshide, Wuhan, China) in a 900 W oven for 5 min to induce epitope retrieval. Slides were allowed to cool at room temperature for 30 min. A primary mouse antibody specific to human PSCA (Boshide, Wuhan, China) with a 1:100 dilution was applied to incubate with the slides at room temperature for 2 h. Labeling was detected by sequentially adding biotinylated secondary antibodies and strepavidin-peroxidase, and localized using 3,3'-diaminobenzidine reaction. Sections were then counterstained with hematoxylin. Substitution of the primary antibody with phosphate-buffered-saline (PBS) served as a negative-staining control.

mRNA in situ hybridization (ISH)

Five-μm-thick tissue sections were deparaffinized and dehydrated, then digested in pepsin solution (4 mg/ml in 3% citric acid) for 20 min at 37.5°C, and further processed for ISH. Digoxigenin-labeled sense and antisense human PSCA RNA probes (obtained from Boshide, Wuhan, China) were hybridized to the sections at 48°C overnight. The posthybridization wash with a high stringency was performed sequentially at 37°C in 2 × standard saline citrate (SSC) for 10 min, in 0.5 × SSC for 15 min and in 0.2 × SSC for 30 min. The slides were then incubated to biotinylated mouse anti-digoxigenin antibody at 37.5°C for 1 h followed by washing in 1 × PBS for 20 min at room temperature, and then to strepavidin-peroxidase at 37.5°C for 20 min followed by washing in 1 × PBS for 15 min at room temperature. Subsequently, the slides were developed with diaminobenzidine and then coun-

Table 1: Correlation of PSCA expression with Gleason score

	Intensity × frequency	
Gleason score	0–6 (%)	9 (%)
2–4	5 (83)	1 (17)
5–7	19 (79)	5 (21)
8–10	5 (28)	13 (72)

Table 2: Correlation of PSCA expression with clinical stage

	Intensity × frequency	
Tumor stage	0–6 (%)	9 (%)
≤B	27 (67.5)	13 (32.5)
≥C	2 (25)	6 (75)

terstained with hematoxylin to localize the hybridization signals. Sections hybridized with the sense control probes routinely did not show any specific hybridization signal above background. All slides were hybridized with PBS to substitute for the probes as a negative control.

Scoring methods

To determine the correlation between the results of PSCA immunostaining and mRNA *in situ* hybridization, the same scoring manners are taken in the present study for PSCA protein staining by IHC and PSCA mRNA staining by ISH. Each slide was read and scored by two independently experienced urological pathologists using Olympus BX-41 light microscopes. The evaluation was done in a blinded fashion. For each section, five areas of similar grade were analyzed semiquantitatively for the fraction of cells staining. Fifty percent of specimens were randomly chosen and rescored to determine the degree of interobserver and intraobserver concordance. There was greater than 95% intra- and interobserver agreement.

The intensity of PSCA expression evaluated microscopically was graded on a scale of 0 to 3+ with 3 being the highest expression observed (0, no staining; 1+, mildly intense; 2+, moderately intense; 3+, severely intense). The staining density was quantified as the percentage of cells staining positive for PSCA with the primary antibody or hybridization probe, as follows: 0 = no staining; 1 = positive staining in <25% of the sample; 2 = positive staining in 25%–50% of the sample; 3 = positive staining in >50% of the sample. Intensity score (0 to 3+) was multiplied by the density score (0–3) to give an overall score of 0–9 [1,5]. In this way, we were able to differentiate specimens that may have had focal areas of increased staining from those that had diffuse areas of increased staining [6]. The overall score for each specimen was then categorically assigned to one of the following groups: 0 score, negative expression; 1–2 scores, weak expression; 3–6 scores, moderate expression; 9 score, strong expression.

Statistical analysis

Intensity and density of PSCA protein and mRNA expression in BPH, PIN and Pca tissues were compared using the Chi-square and Student's *t*-test. Univariate associations between PSCA expression and Gleason score, clinical stage and progression to androgen-independence were calculated using Fisher's Exact Test. For all analyses, $p < 0.05$ was considered statistically significant.

Results

PSCA expression in BPH

In general, PSCA protein and mRNA were expressed weakly in individual samples of BPH. Some areas of prostate expressed weak levels (composite score 1–2), whereas other areas were completely negative (composite score 0). Four cases (20%) of BPH had moderate expression of PSCA protein and mRNA (composite score 4–6) by IHC and ISH. In 2/20 (10%) BPH specimens, PSCA mRNA expression was moderate (composite score 3–6), but PSCA protein expression was weak (composite score

2) in one and negative (composite score 0) in the other. PSCA expression was localized to the basal and secretory epithelial cells, and prostatic stroma was almost negative staining for PSCA protein and mRNA in all cases examined.

PSCA expression in PIN

In this study, we detected weak or negative expression of PSCA protein and mRNA (≤2 scores) in 7 of 9 (77.8%) low grade PIN and in 2 of 11 (18.2%) HGPIN, and moderate expression (3–6 scores) in the rest 2 low grade PIN and 5 of 11 (45.5%) HGPIN. One HGPIN with moderate PSCA mRNA expression (6 score) was found weak staining for PSCA protein (2 score) by IHC. Strong PSCA protein and mRNA expression (9 score) were detected in the remaining 3 of 11 (27.3%) HGPIN. There was a statistically significant difference of PSCA protein and mRNA expression levels observed between HGPIN and BPH ($p < 0.05$), but no statistical difference reached between low grade PIN and BPH ($p > 0.05$).

PSCA expression in Pca

In order to determine if PSCA protein and mRNA can be detected in prostate cancers and if PSCA expression levels are increased in malignant compared with benign glands, Forty-eight paraffin-embedded Pca specimens were analysed by IHC and ISH. It was shown that 19 of 48 (39.6%) Pca samples stained very strongly for PSCA protein and mRNA with a score of 9 and another 21 (43.8%) specimens displayed moderate staining with scores of 4–6 (Figure 1). In addition, 4 specimens with moderate to strong PSCA mRNA expression (scores of 4–9) had weak protein staining (a score of 2) by IHC analyses. Overall, Pca expressed a significantly higher level of PSCA protein and mRNA than any other specimen category in this study ($p < 0.05$, compared with BPH and PIN respectively). The result demonstrates that PSCA protein and mRNA are overexpressed by a majority of human Pca.

Correlation of PSCA expression with Gleason score in Pca

Using the semi-quantitative scoring method as described in Materials and Methods, we compared the expression level of PSCA protein and mRNA with Gleason grade of Pca, as shown in Table 1. Prostate adenocarcinomas were graded by Gleason score as 2–4 scores = well-differentiation, 5–7 scores = moderate-differentiation and 8–10 scores = poor-differentiation [7]. Seventy-two percent of Gleason scores 8–10 prostate cancers had very strong staining of PSCA compared to 21% with Gleason scores 5–7 and 17% with 2–4 respectively, demonstrating that poorly differentiated Pca had significantly stronger expression of PSCA protein and mRNA than moderately and well differentiated tumors ($p < 0.05$). As depicted in Figure 1, IHC and ISH analyses showed that PSCA protein and mRNA expression in several cases of poorly differentiated Pca were particularly prominent, with more intense and uniform staining. The results indicate that PSCA expression increases significantly with higher tumor grade in human Pca.

Correlation of PSCA expression with clinical stage in Pca

With regards to PSCA expression in every stage of Pca, we showed the results in Table 2. Seventy-five percent of locally advanced and node positive cancers (i.e. C-D stages) expressed statistically high levels of PSCA versus 32.5% that were organ confined (i.e. A-B stages) ($p < 0.05$). The data demonstrate that PSCA expression increases significantly with advanced tumor stage in human Pca.

Correlation of PSCA expression with androgen-independent progression of Pca

All 9 specimens of androgen-independent prostate cancers stained positive for PSCA protein and mRNA. Eight specimens were obtained from patients managed prior to androgen ablation therapy. Seven of eight (87.5%) of these androgen-independent prostate cancers were in the strongest staining category (score = 9), compared with three out of eight (37.5%) of patients with androgen-dependent cancers ($p < 0.05$). The results demonstrate that PSCA expression increases significantly with progression to androgen-independence of human Pca.

It is evident from the results above that within a majority of human prostate cancers the level of PSCA protein and mRNA expression correlates significantly with increasing grade, worsening stage and progression to androgen-independence.

Correlation of PSCA immunostaining and mRNA in situ hybridization

In all 88 specimens surveyed herein, we compared the results of PSCA IHC staining with mRNA ISH analysis. Positive staining areas and its intensity and density scores evaluated by IHC were identical to those seen by ISH in 79 of 88 (89.8%) specimens (18/20 BPH, 19/20 PIN and 42/48 Pca respectively). Importantly, 27/27 samples with PSCA mRNA composite scores of 0–2, 32/36 samples with scores of 3–6 and 22/24 samples with a score of 9 also had PSCA protein expression scores of 0–2, 3–6 and 9 respectively. However, in 5 samples with PSCA mRNA overall scores of 3–6 and in 2 with scores of 9 there were less or negative PSCA protein expression (i.e. scores of 0–4), suggesting that this may reflect posttranscriptional modification of PSCA or that the epitopes recognized by PSCA mAb may be obscured in some cancers. The data demonstrate that the results of PSCA immunostaining were consistent with those of mRNA ISH analysis, showing a high degree of correlation between PSCA protein and mRNA expression.

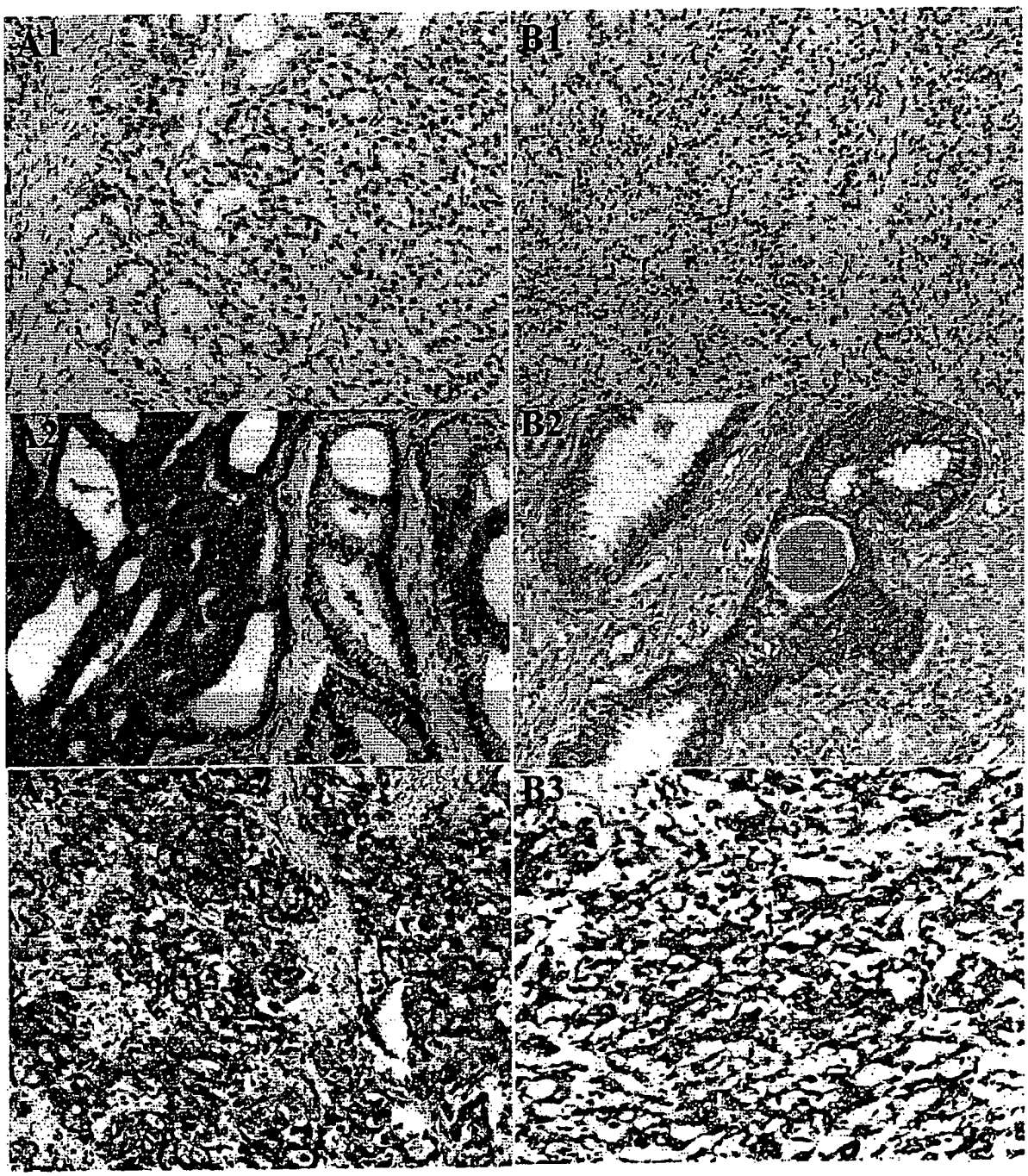


Figure 1
Representatives of PSCA IHC and ISH staining in Pca (A. IHC staining, B. ISH staining, ×200 magnification). A_1, B_1: negative control of IHC and ISH. PBS replacing the primary antibody (A_1) and hybridization with a sense PSCA probe (B_1) showed no background staining. A_2, B_2: a moderately differentiated Pca (Gleason score = 3+3 = 6) with moderate staining (composite score = 6) in all malignant cells; A_2: IHC shows not only cell surface but also apparent cytoplasmic staining of PSCA protein. A_3, B_3: a poorly differentiated Pca (Gleason score = 4+4 = 8) with very strong staining (composite score = 9) in all malignant cells.

Discussion

PSCA is homologous to a group of cell surface proteins that mark the earliest phase of hematopoietic development. PSCA mRNA expression is prostate-specific in normal male tissues and is highly up-regulated in both androgen-dependent and-independent Pca xenografts (LAPC-4 tumors). We hypothesize that PSCA may play a role in Pca tumorigenesis and progression, and may serve as a target for Pca diagnosis and treatment. In this study, IHC and ISH showed that in general there were weak or absent PSCA protein and mRNA expression in BPH and low grade PIN tissues. However, PSCA protein and mRNA are widely expressed in HGPIN, the putative precursor of invasive Pca, suggesting that up-regulation of PSCA is an early event in prostate carcinogenesis. Recently, Reiter RE et al [1], using ISH analysis, reported that 97 of 118 (82%) HGPIN specimens stained strongly positive for PSCA mRNA. A very similar finding was seen on mouse PSCA (mPSCA) expression in mouse HGPIN tissues by Tran C. P et al [8]. These data suggest that PSCA may be a new marker associated with transformation of prostate cells and tumorigenesis.

We observed that PSCA protein and mRNA are highly expressed in a large percentage of human prostate cancers, including advanced, poorly differentiated, androgen-independent and metastatic cases. Fluorescence-activated cell sorting and confocal/ immunofluorescent studies demonstrated cell surface expression of PSCA protein in Pca cells [9]. Our IHC expression analysis of PSCA shows not only cell surface but also apparent cytoplasmic staining of PSCA protein in Pca specimens (Figure 1). One possible explanation for this is that anti-PSCA antibody can recognize PSCA peptide precursors that reside in the cytoplasm. Also, it is possible that the positive staining that appears in the cytoplasm is actually from the overlying cell membrane [5]. These data seem to indicate that PSCA is a novel cell surface marker for human Pca.

Our results show that elevated level of PSCA expression correlates with high grade (i.e. poor differentiation), increased tumor stage and progression to androgen-independence of Pca. These findings support the original IHC analyses by Gu Z et al [9], who reported that PSCA protein expressed in 94% of primary Pca and the intensity of PSCA protein expression increased with tumor grade, stage and progression to androgen-independence. Our results also collaborate the recent work of Han KR et al [10], in which the significant association between high PSCA expression and adverse prognostic features such as high Gleason score, seminal vesicle invasion and capsular involvement in Pca was found. It is suggested that PSCA overexpression may be an adverse predictor for recurrence, clinical progression or survival of Pca. Hara H et al [11] used RT-PCR detection of PSA, PSMA and PSCA in 1 ml of peripheral blood to evaluate Pca patients with poor prognosis. The results showed that among 58 PCa patients, each PCR indicated the prognostic value in the hierarchy of PSCA>PSA>PSMA RT-PCR, and extraprostatic cases with positive PSCA PCR indicated lower disease-progression-free survival than those with negative PSCA PCR, demonstrating that PSCA can be used as a prognostic factor. Dubey P et al [12] reported that elevated numbers of PSCA + cells correlate positively with the onset and development of prostate carcinoma over a long time span in the prostates of the TRAMP and PTEN +/- models compared with its normal prostates. Taken together with our present findings, in which PSCA is overexpressed from HGPIN to almost frank carcinoma, it is reasonable and possible to use increased PSCA expression level or increased numbers of PSCA-positive cells in the prostate samples as a prognostic marker to predict the potential onset of this cancer. These data raise the possibility that PSCA may have diagnostic utility or clinical prognostic value in human Pca.

The cause of PSCA overexpression in Pca is not known. One possible mechanism is that it may result from PSCA gene amplification. In humans, PSCA is located on chromosome 8q24.2 [1], which is often amplified in metastatic and recurrent Pca and considered to indicate a poor prognosis [13-15]. Interestingly, PSCA is in close proximity to the c-myc oncogene, which is amplified in >20% of recurrent and metastatic prostate cancers [16,17]. Reiter RE et al [18] reported that PSCA and MYC gene copy numbers were co-amplified in 25% of tumors (five out of twenty), demonstrating that PSCA overexpression is associated with PSCA and MYC coamplification in Pca. Gu Z et al [9] recently reporteted that in 102 specimens available to compare the results of PSCA immunostaining with their previous mRNA ISH analysis, 92 (90.2%) had identically positive areas of PSCA protein and mRNA expression. Taken together with our findings, in which we detected moderate to strong expression of PSCA protein and mRNA in 34 of 40 (85%) Pca specimens examined simultaneously by IHC and ISH analyses, it is demonstrated that PSCA protein and mRNA overexpressed in human Pca, and that the increased protein level of PSCA was resulted from the upregulated transcription of its mRNA.

At present, the regulation mechanisms of human PSCA expression and its biological function are yet to be elucidated. PSCA expression may be regulated by multiple factors [18]. Watabe T et al [19] reported that transcriptional control is a major component regulating PSCA expression levels. In addition, induction of PSCA expression may be regulated or mediated through cell-cell contact and protein kinase C (PKC) [20]. Homologues of PSCA have diverse activities, and have themselves been involved in

carcinogenesis. Signalling through SCA-2 has been demonstrated to prevent apoptosis in immature thymocytes [21]. Thy-1 is involved in T cell activation and transducts signals through src-like tyrosine kinases [22]. Ly-6 genes have been implicated both in tumorigenesis and in cell-cell adhesion [23-25]. Cell-cell or cell-matrix interaction is critical for local tumor growth and spread to distal sites. From its restricted expression in basal cells of normal prostate and its homology to SCA-2, PSCA may play a role in stem/progenitor cell function, such as self-renewal (i.e. anti-apoptosis) and/or proliferation [1]. Taken together with the results in the present study, we speculate that PSCA may play a role in tumorigenesis and clinical progression of Pca through affecting cell transformation and proliferation. From our results, it is also suggested that PSCA as a new cell surface antigen may have a number of potential uses in the diagnosis, therapy and clinical prognosis of human Pca. PSCA overexpression in prostate biopsies could be used to identify patients at high risk to develop recurrent or metastatic disease, and to discriminate cancers from normal glands in prostatectomy samples. Similarly, the detection of PSCA-overexpressing cells in bone marrow or peripheral blood may identify and predict metastatic progression better than current assays, which identify only PSA-positive or PSMA-positive prostate cells.

In summary, we have shown in this study that PSCA protein and mRNA are maintained in expression from HGPIN through all stages of Pca in a majority of cases, which may be associated with prostate carcinogenesis and correlate positively with high tumor grade (poor cell differentiation), advanced stage and androgen-independent progression. PSCA protein overexpression is due to the upregulation of its mRNA transcription. The results suggest that PSCA may be a promising molecular marker for the clinical prognosis of human Pca and a valuable target for diagnosis and therapy of this tumor.

Competing interests

None declared.

References

1. Reiter RE, Gu Z, Watabe T, Thomas G, Szigeti K, David E, Wahl M, Nisitani S, Yamashiro J, Le Beau MM, Loda M, Witte ON: **Prostate stem cell antigen: a cell surface marker overexpressed in prostate cancer.** *Proc Natl Acad Sci USA* 1998, 95:1735-1740.
2. Gleason DF: **Histologic grading and clinical staging of prostatic carcinoma.** In: *Urologic Pathology: The Prostate* Edited by: Tannebaum M. Philadelphia, Lea & Febiger; 1977:171-197.
3. Brawer MK: **Prostatic intraepithelial neoplasia: a premalignant lesion.** *Hum Pathol* 1992, 23:242-248.
4. Amin MB, Ro JY, Ayala AC: **Prostatic intraepithelial neoplasia: relationship to adenocarcinoma of prostate.** *Pathol Annu* 1994, 29:1-30.
5. Amara N, Palapattu GS, Schrage M, Gu Z, Thomas GV, Dorey F, Said J, Reiter RE: **Prostate stem cell antigen is overexpressed in human transitional cell carcinoma.** *Cancer Res* 2001, 61:4660-4665.
6. Hanas JS, Lerner MR, Lightfoot SA, Raczkowski C, Kastens DJ, Brackett DJ, Postier RG: **Expression of the cyclin-dependent kinase inhibitor p21 (WAF1/CIP1) and p53 tumor suppressor in dysplastic progression and adenocarcinoma in Barrett esophagus.** *Cancer (Phila)* 1999, 86:756-763.
7. Egevad L, Gramfors T, Karlberg L: **Prognostic value of the Gleason score in prostate cancer.** *BJU Int* 2002, 89:538-542.
8. Tran CP, Lin C, Yamashiro J, Reiter RE: **Prostate stem cell antigen is a marker of late intermediate prostate epithelial cells.** *Mol Cancer Res* 2002, 1:113-121.
9. Gu Z, Thomas G, Yamashiro J, Shintaku IP, Dorey F, Raitano A, Witte ON, Said JW, Loda M, Reiter RE: **Prostate stem cell antigen (PSCA) expression increases with high Gleason score, advanced stage and bone metastasis in prostate cancer.** *Oncogene* 2000, 19:1288-1296.
10. Han KR, Seligson DB, Liu X, Horvath S, Shintaku PI, Thomas GV, Said JW, Reiter RE: **Prostate stem cell antigen expression is associated with gleason score, seminal vesicle invasion and capsular invasion in prostate cancer.** *J Urol* 2004, 171:1117-1121.
11. Hara H, Kasahara T, Kawasaki T, Bilim V, Obara K, Takahashi K, Tomita Y: **Reverse Transcription-Polymerase Chain Reaction Detection of Prostate-specific Antigen, Prostate-specific Membrane Antigen, and Prostate Stem Cell Antigen in One Milliliter of Peripheral Blood.** *Clin Cancer Res* 2002, 8:1794-1799.
12. Dubey P, Wu H, Reiter RE, Witte ON: **Alternative pathways to prostate carcinoma activate prostate stem cell antigen expression.** *Cancer Res* 2001, 61:3256-3261.
13. Visa korpi T, Kallioniemi AH, Syvanen AC, Hyytinen ER, Karhu R, Tammela T, Isola JJ, Kallioniemi OP: **Genetic changes in primary and recrrent prostate cancer by comparative genomic hybridization.** *Cancer Res* 1995, 55:342-347.
14. Sato K, Qian J, Slezak JM, Lieber MM, Bostwick DG, Bergstralh EJ, Jenkins RB: **Clinical significance of alterations of chromosome 8 in high-grade, advanced, nonmetastatic prostate carcinoma.** *J Natl Cancer Inst* 1999, 91:1574-1580.
15. Van Den Berg C, Guan XY, Von Hoff D, Jenkins R, Bittner , Griffin C, Kallioniemi O, Visakorpi , McGill , Herath J, Epstein J, Sarosdy M, Meltzer P, Trent J: **DNA sequence amplification in human prostate cancer identified by chromosome microdissection: potential prognostic implications.** *Clin Cancer Res* 1995, 1:11-18.
16. Jenkins RB, Qian J, Lieber MM, Bostwick DG: **Detection of c-myc oncogene amplification and chromosomal anomalies in metastatic prostatic carcinoma by fluorescence in situ hybridization.** *Cancer Res* 1997, 57:524-531.
17. Nupponen NN, Kakkola L, Koivisto P, Visakorpi T: **Genetic alterations in hormone-refractory recurrent prostate carcinomas.** *Am J Pathol* 1998, 153:141-148.
18. Reiter RE, Sato I, Thomas G, Qian J, Gu Z, Watabe T, Loda M, Jenkins RB: **Coamplification of prostate stem cell antigen (PSCA) and MYC in locally advanced prostate cancer.** *Genes Chromosomes Cancer* 2000, 27:95-103.
19. Watabe T, Lin M, Donjacour AA, Cunha GR, Witte ON, Reiter RE: **Growth, regeneration, and tumorigenesis of the prostate activates the PSCA promoter.** *Proc Natl Acad Sci USA* 2002, 99:401-406.
20. Bahrenberg G, Brauers A, Joost HG, Jakse G: **PSCA expression is regulated by phorbol ester and cell adhesion in the bladder carcinoma cell line RT112.** *Cancer Lett* 2001, 168:37-43.
21. Noda S, Kosugi A, Saitoh S, Narumiya S, Hamaoka T: **Protection from anti-TCR/CD3-induced apoptosis in immature thymocytes by a signal through thymic shared antigen-1/stem cell antigen-2.** *J Exp Med* 1996, 183:2355-2360.
22. Thomas PM, Samelson LE: **The glycophosphatidylinositol-anchored Thy-1 molecule interacts with the p60fyn protein tyrosine kinase in T cells.** *J Biol Chem* 1992, 267:12317-12322.
23. Bamezai A, Rock KL: **Overexpressed Ly-6A.2 mediated cell-cell adhesion by binding a ligand expressed on lymphoid cells.** *Proc Natl Acad Sci USA* 1995, 92:4294-4298.
24. Katz BZ, Eshel R, Sagi-Assif O, Witz IP: **An association between high Ly-6A/E expression on tumor cells and a highly malignant phenotype.** *Int J Cancer* 1994, 59:684-691.
25. Brakenhoff RH, Gerretsen M, Knippels EM, van Dijk M, van Essen H, Weghuis DO, Sinke RJ, Snow GB, van Dongen GA: **The human E48 antigen, highly homologous to the murine Ly-6 antigen ThB, is a GPI-anchored molecule apparently involved in keratinocyte cell-cell adhesion.** *J Cell Biol* 1995, 129:1677-1689.

EXHIBIT 6

Review

Translation Initiation in Cancer: A Novel Target for Therapy[1]

Funda Meric[2] and Kelly K. Hunt

Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030

Abstract

Translation initiation is regulated in response to nutrient availability and mitogenic stimulation and is coupled with cell cycle progression and cell growth. Several alterations in translational control occur in cancer. Variant mRNA sequences can alter the translational efficiency of individual mRNA molecules, which in turn play a role in cancer biology. Changes in the expression or availability of components of the translational machinery and in the activation of translation through signal transduction pathways can lead to more global changes, such as an increase in the overall rate of protein synthesis and translational activation of the mRNA molecules involved in cell growth and proliferation. We review the basic principles of translational control, the alterations encountered in cancer, and selected therapies targeting translation initiation to help elucidate new therapeutic avenues.

Introduction

The fundamental principle of molecular therapeutics in cancer is to exploit the differences in gene expression between cancer cells and normal cells. With the advent of cDNA array technology, most efforts have concentrated on identifying differences in gene expression at the level of mRNA, which can be attributable either to DNA amplification or to differences in transcription. Gene expression is quite complicated, however, and is also regulated at the level of mRNA stability, mRNA translation, and protein stability.

The power of translational regulation has been best recognized among developmental biologists, because transcription does not occur in early embryogenesis in eukaryotes. For example, in *Xenopus*, the period of transcriptional quiescence continues until the embryo reaches midblastula transition, the 4000-cell stage. Therefore, all necessary mRNA molecules are transcribed during oogenesis and stockpiled in a translationally inactive, masked form. The mRNA are translationally activated at appropriate times during oocyte maturation, fertilization, and early embryogenesis and thus, are under strict translational control.

Translation has an established role in cell growth. Basically, an increase in protein synthesis occurs as a consequence of mitogenesis. Until recently, however, little was known about the alterations in mRNA translation in cancer, and much is yet to be discovered about their role in the development and progression of cancer. Here we review the basic principles of translational control, the alterations encountered in cancer, and selected therapies targeting translation initiation to elucidate potential new therapeutic avenues.

Basic Principles of Translational Control

Mechanism of Translation Initiation

Translation initiation is the main step in translational regulation. Translation initiation is a complex process in which the initiator tRNA and the 40S and 60S ribosomal subunits are recruited to the 5′ end of a mRNA molecule and assembled by eukaryotic translation initiation factors into an 80S ribosome at the start codon of the mRNA (Fig. 1). The 5′ end of eukaryotic mRNA is capped, *i.e.*, contains the cap structure m^7GpppN (7-methylguanosine-triphospho-5′-ribonucleoside). Most translation in eukaryotes occurs in a cap-dependent fashion, *i.e.*, the cap is specifically recognized by the eIF4E,[3] which binds the 5′ cap. The eIF4F translation initiation complex is then formed by the assembly of eIF4E, the RNA helicase eIF4A, and eIF4G, a scaffolding protein that mediates the binding of the 40S ribosomal subunit to the mRNA molecule through interaction with the eIF3 protein present on the 40S ribosome. eIF4A and eIF4B participate in melting the secondary structure of the 5′ UTR of the mRNA. The 43S initiation complex (40S/eIF2/Met-tRNA/GTP complex) scans the mRNA in a 5′→3′ direction until it encounters an AUG start codon. This start codon is then base-paired to the anticodon of initiator tRNA, forming the 48S initiation complex. The initiation factors are then displaced from the 48S complex, and the 60S ribosome joins to form the 80S ribosome.

Unlike most eukaryotic translation, translation initiation of certain mRNAs, such as the picornavirus RNA, is cap independent and occurs by internal ribosome entry. This mechanism does not require eIF4E. Either the 43S complex can bind the initiation codon directly through interaction with the IRES in the 5′ UTR such as in the encephalomyocarditis virus, or it can

Received 5/16/02; revised 7/12/02; accepted 7/22/02.

[1] F. M. is supported by The University of Texas M. D. Anderson Cancer Center Physician-Scientist Program and by NIH Grant 1K08-CA 91895-01. K. K. H. is supported by Department of Defense Award DAMD-17-97-1-7162.

[2] To whom requests for reprints should be addressed, at Department of Surgical Oncology, Box 444, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: (713) 745-4453; Fax: (713) 745-4926; E-mail: fmeric@mdanderson.org.

[3] The abbreviations used are: eIF4E, eukaryotic initiation factor 4E; UTR, untranslated region; IRES, internal ribosome entry site; 4E-BP1, eukaryotic initiation factor 4E-binding protein 1; S6K, ribosomal p70 S6 kinase; mTOR, mammalian target of rapamycin; ATM, ataxia telangiectasia mutated; PI3K, phosphatidylinositol 3-kinase; PTEN, phosphatase and tensin homolog deleted from chromosome 10; PP2A, protein phosphatase 2A; TGF-β3, transforming growth factor-β3; PAP, poly(A) polymerase; EPA, eicosopentaenoic acid; *mda-7*, melanoma differentiation-associated gene 7.

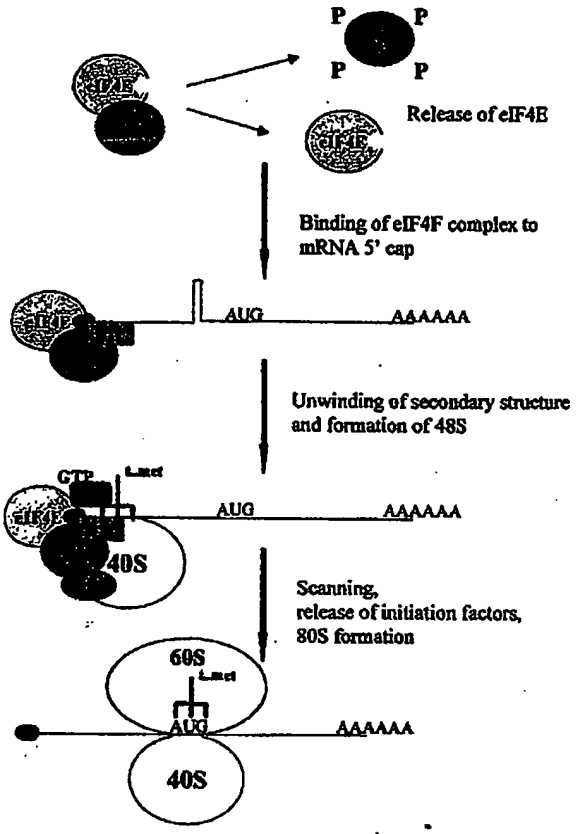


Fig. 1. Translation initiation in eukaryotes. The 4E-BPs are hyperphosphorylated to release eIF4E so that it can interact with the 5′ cap, and the eIF4F initiation complex is assembled. The interaction of poly(A) binding protein with the initiation complex and circularization of the mRNA is not depicted in the diagram. The secondary structure of the 5′ UTR is melted, the 40S ribosomal subunit is bound to eIF3, and the ternary complex consisting of eIF2, GTP, and the Met-tRNA are recruited to the mRNA. The ribosome scans the mRNA in a 5′→3′ direction until an AUG start codon is found in the appropriate sequence context. The initiation factors are released, and the large ribosomal subunit is recruited.

initially attach to the IRES and then reach the initiation codon by scanning or transfer, as is the case with the poliovirus (1).

Regulation of Translation Initiation

Translation initiation can be regulated by alterations in the expression or phosphorylation status of the various factors involved. Key components in translational regulation that may provide potential therapeutic targets follow.

eIF4E. eIF4E plays a central role in translation regulation. It is the least abundant of the initiation factors and is considered the rate-limiting component for initiation of cap-dependent translation. eIF4E may also be involved in mRNA splicing, mRNA 3′ processing, and mRNA nucleocytoplasmic transport (2). eIF4E expression can be increased at the transcriptional level in response to serum or growth factors (3). eIF4E overexpression may cause preferential translation of mRNAs containing excessive secondary structure in their 5′ UTR that are normally discriminated against by the translational machinery and thus are inefficiently translated (4–7). As examples of this, overexpression of eIF4E promotes increased translation of vascular endothelial growth factor, fibroblast growth factor-2, and cyclin D1 (2, 8, 9).

Another mechanism of control is the regulation of eIF4E phosphorylation. eIF4E phosphorylation is mediated by the mitogen-activated protein kinase-interacting kinase 1, which is activated by the mitogen-activated pathway activating extracellular signal-related kinases and the stress-activated pathway acting through p38 mitogen-activated protein kinase (10–13). Several mitogens, such as serum, platelet-derived growth factor, epidermal growth factor, insulin, angiotensin II, src kinase overexpression, and ras overexpression, lead to eIF4E phosphorylation (14). The phosphorylation status of eIF4E is usually correlated with the translational rate and growth status of the cell; however, eIF4E phosphorylation has also been observed in response to some cellular stresses when translational rates actually decrease (15). Thus, further study is needed to understand the effects of eIF4E phosphorylation on eIF4E activity.

Another mechanism of regulation is the alteration of eIF4E availability by the binding of eIF4E to the eIF4E-binding proteins (4E-BP, also known as PHAS-I). 4E-BPs compete with eIF4G for a binding site in eIF4E. The binding of eIF4E to the best characterized eIF4E-binding protein, 4E-BP1, is regulated by 4E-BP1 phosphorylation. Hypophosphorylated 4E-BP1 binds to eIF4E, whereas 4E-BP1 hyperphosphorylation decreases this binding. Insulin, angiotensin, epidermal growth factor, platelet-derived growth factor, hepatocyte growth factor, nerve growth factor, insulin-like growth factors I and II, interleukin 3, granulocyte-macrophage colony-stimulating factor + steel factor, gastrin, and the adenovirus have all been reported to induce phosphorylation of 4E-BP1 and to decrease the ability of 4E-BP1 to bind eIF4E (15, 16). Conversely, deprivation of nutrients or growth factors results in 4E-BP1 dephosphorylation, an increase in eIF4E binding, and a decrease in cap-dependent translation.

p70 S6 Kinase. Phosphorylation of ribosomal 40S protein S6 by S6K is thought to play an important role in translational regulation. S6K −/− mouse embryonic cells proliferate more slowly than do parental cells, demonstrating that S6K has a positive influence on cell proliferation (17). S6K regulates the translation of a group of mRNAs possessing a 5′ terminal oligopyrimidine tract (5′ TOP) found at the 5′ UTR of ribosomal protein mRNAs and other mRNAs coding for components of the translational machinery. Phosphorylation of S6K is regulated in part based on the availability of nutrients (18, 19) and is stimulated by several growth factors, such as platelet-derived growth factor and insulin-like growth factor I (20).

eIF2α Phosphorylation. The binding of the initiator tRNA to the small ribosomal unit is mediated by translation initiation factor eIF2. Phosphorylation of the α-subunit of eIF2 prevents formation of the eIF2/GTP/Met-tRNA complex and inhibits global protein synthesis (21, 22). eIF2α is phosphorylated under a variety of conditions, such as viral infection, nutrient deprivation, heme deprivation, and apoptosis (22). eIF2α is phosphorylated by heme-regulated inhibitor, nutrient-regulated protein kinase, and the IFN-induced, double-stranded RNA-activated protein kinase (PKR; Ref. 23).

The mTOR Signaling Pathway. The macrolide antibiotic rapamycin (Sirolimus; Wyeth-Ayerst Research, Collegeville, PA) has been the subject of intensive study because it inhibits signal transduction pathways involved in T-cell activation. The rapamycin-sensitive component of these pathways is mTOR (also called FRAP or RAFT1). mTOR is the mammalian homologue of the yeast TOR proteins that regulate G_1 progression and translation in response to nutrient availability (24). mTOR is a serine-threonine kinase that modulates translation initiation by altering the phosphorylation status of 4E-BP1 and S6K (Fig. 2; Ref. 25).

4E-BP1 is phosphorylated on multiple residues. mTOR phosphorylates the Thr-37 and Thr-46 residues of 4E-BP1 *in vitro* (26); however, phosphorylation at these sites is not associated with a loss of eIF4E binding. Phosphorylation of Thr-37 and Thr-46 is required for subsequent phosphorylation at several COOH-terminal, serum-sensitive sites; a combination of these phosphorylation events appears to be needed to inhibit the binding of 4E-BP1 to eIF4E (25). The product of the *ATM* gene, p38/MSK1 pathway, and protein kinase Cσ also play a role in 4E-BP1 phosphorylation (27–29).

S6K and 4E-BP1 are also regulated, in part, by PI3K and its downstream protein kinase Akt. PTEN is a phosphatase that negatively regulates PI3K signaling. PTEN null cells have constitutively active of Akt, with increased S6K activity and S6 phosphorylation (30). S6K activity is inhibited both by PI3K inhibitors wortmannin and LY294002 and by mTOR inhibitor rapamycin (24). Akt phosphorylates Ser-2448 in mTOR *in vitro*, and this site is phosphorylated upon Akt activation *in vivo* (31–33). Thus, mTOR is regulated by the PI3K/Akt pathway; however, this does not appear to be the only mode of regulation of mTOR activity. Whether the PI3K pathway also regulates S6K and 4E-BP1 phosphorylation independent of mTOR is controversial.

Interestingly, mTOR autophosphorylation is blocked by wortmannin but not by rapamycin (34). This seeming inconsistency suggests that mTOR-responsive regulation of 4E-BP1 and S6K activity occurs through a mechanism other than intrinsic mTOR kinase activity. An alternate pathway for 4E-BP1 and S6K phosphorylation by mTOR activity is by the inhibition of a phosphatase. Treatment with calyculin A, an inhibitor of phosphatases 1 and 2A, reduces rapamycin-induced dephosphorylation of 4E-BP1 and S6K by rapamycin (35). PP2A interacts with full-length S6K but not with a S6K mutant that is resistant to dephosphorylation resulting from rapamycin. mTOR phosphorylates PP2A *in vitro*; however, how this process alters PP2A activity is not known. These results are consistent with the model that phosphorylation of a phosphatase by mTOR prevents dephosphorylation of 4E-BP1 and S6K, and conversely, that nutrient deprivation and rapamycin block inhibition of the phosphatase by mTOR.

Polyadenylation. The poly(A) tail in eukaryotic mRNA is important in enhancing translation initiation and mRNA stability. Polyadenylation plays a key role in regulating gene expression during oogenesis and early embryogenesis. Some mRNA that are translationally inactive in the oocyte are polyadenylated concomitantly with translational activation in oocyte maturation, whereas other mRNAs that are translationally active during oogenesis are deadenylated and translationally silenced (36–38). Thus, control of poly(A) tail synthesis is an important regulatory step in gene expression. The 5′ cap and poly(A) tail are thought to function synergistically to regulate mRNA translational efficiency (39, 40).

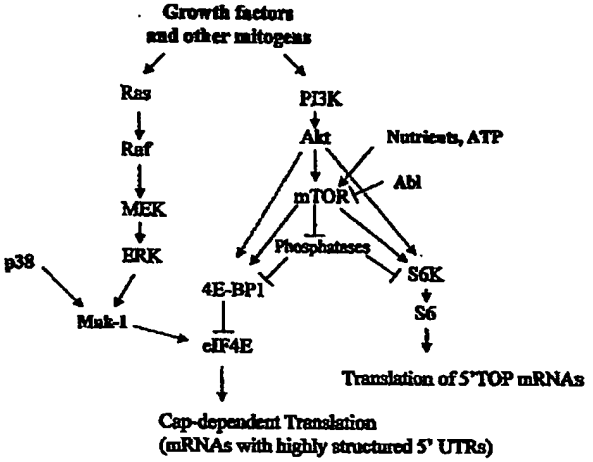


Fig. 2. Regulation of translation initiation by signal transduction pathways. Signaling via p38, extracellular signal-related kinase, PI3K, and mTOR can all activate translation initiation.

RNA Packaging. Most RNA-binding proteins are assembled on a transcript at the time of transcription, thus determining the translational fate of the transcript (41). A highly conserved family of Y-box proteins is found in cytoplasmic messenger ribonucleoprotein particles, where the proteins are thought to play a role in restricting the recruitment of mRNA to the translational machinery (41–43). The major mRNA-associated protein, YB-1, destabilizes the interaction of eIF4E and the 5′ mRNA cap *in vitro*, and overexpression of YB-1 results in translational repression *in vivo* (44). Thus, alterations in RNA packaging can also play an important role in translational regulation.

Translation Alterations Encountered in Cancer

Three main alterations at the translational level occur in cancer: variations in mRNA sequences that increase or decrease translational efficiency, changes in the expression or availability of components of the translational machinery, and activation of translation through aberrantly activated signal transduction pathways. The first alteration affects the translation of an individual mRNA that may play a role in carcinogenesis. The second and third alterations can lead to more global changes, such as an increase in the overall rate of protein synthesis, and the translational activation of several mRNA species.

Variations in mRNA Sequence

Variations in mRNA sequence affect the translational efficiency of the transcript. A brief description of these variations and examples of each mechanism follow.

Mutations. Mutations in the mRNA sequence, especially in the 5′ UTR, can alter its translational efficiency, as seen in the following examples.

c-myc. Saito *et al.* proposed that translation of full-length *c-myc* is repressed, whereas in several Burkitt lymphomas that have deletions of the mRNA 5′ UTR, translation of *c-myc* is more efficient (45). More recently, it was reported that the 5′ UTR of *c-myc* contains an IRES, and thus *c-myc* translation can be initiated by a cap-independent as well as a cap-dependent mechanism (46, 47). In patients with multiple myeloma, a C→T mutation in the c-myc IRES was identified (48) and found to cause an enhanced initiation of translation via internal ribosomal entry (49).

BRCA1. A somatic point mutation (117 G→C) in position −3 with respect to the start codon of the *BRCA1* gene was identified in a highly aggressive sporadic breast cancer (50). Chimeric constructs consisting of the wild-type or mutated *BRCA1* 5′ UTR and a downstream luciferase reporter demonstrated a decrease in the translational efficiency with the 5′ UTR mutation.

Cyclin-dependent Kinase Inhibitor 2A. Some inherited melanoma kindreds have a G→T transversion at base −34 of cyclin-dependent kinase inhibitor-2A, which encodes a cyclin-dependent kinase 4/cyclin-dependent kinase 6 kinase inhibitor important in G_1 checkpoint regulation (51). This mutation gives rise to a novel AUG translation initiation codon, creating an upstream open reading frame that competes for scanning ribosomes and decreases translation from the wild-type AUG.

Alternate Splicing and Alternate Transcription Start Sites. Alterations in splicing and alternate transcription sites can lead to variations in 5′ UTR sequence, length, and secondary structure, ultimately impacting translational efficiency.

ATM. The *ATM* gene has four noncoding exons in its 5′ UTR that undergo extensive alternative splicing (52). The contents of 12 different 5′ UTRs that show considerable diversity in length and sequence have been identified. These divergent 5′ leader sequences play an important role in the translational regulation of the *ATM* gene.

mdm. In a subset of tumors, overexpression of the oncoprotein mdm2 results in enhanced translation of the *mdm2* mRNA. Use of different promoters leads to two *mdm2* transcripts that differ only in their 5′ leaders (53). The longer 5′ UTR contains two upstream open reading frames, and this mRNA is loaded with ribosomes inefficiently compared with the short 5′ UTR.

BRCA1. In a normal mammary gland, *BRCA1* mRNA is expressed with a shorter leader sequence (5′UTRa), whereas in sporadic breast cancer tissue, *BRCA1* mRNA is expressed with a longer leader sequence (5′ UTRb); the translational efficiency of transcripts containing 5′ UTRb is 10 times lower than that of transcripts containing 5′ UTRa (54).

TGF-β3. *TGF-β3* mRNA includes a 1.1-kb 5′ UTR, which exerts an inhibitory effect on translation. Many human breast cancer cell lines contain a novel *TGF-β3* transcript with a 5′ UTR that is 870 nucleotides shorter and has a 7-fold greater translational efficiency than the normal *TGF-β3* mRNA (55).

Alternate Polyadenylation Sites. Multiple polyadenylation signals leading to the generation of several transcripts with differing 3′ UTR have been described for several mRNA species, such as the *RET* proto-oncogene (56), *ATM* gene (52), tissue inhibitor of metalloproteinases-3 (57), *RHOA* proto-oncogene (58), and calmodulin-I (59). Although the effect of these alternate 3′ UTRs on translation is not yet known, they may be important in RNA-protein interactions that affect translational recruitment. The role of these alterations in cancer development and progression is unknown.

Alterations in the Components of the Translation Machinery

Alterations in the components of translation machinery can take many forms.

Overexpression of eIF4E. Overexpression of eIF4E causes malignant transformation in rodent cells (60) and the deregulation of HeLa cell growth (61). Polunovsky *et al.* (62) found that eIF4E overexpression substitutes for serum and individual growth factors in preserving viability of fibroblasts, which suggests that eIF4E can mediate both proliferative and survival signaling.

Elevated levels of eIF4E mRNA have been found in a broad spectrum of transformed cell lines (63). eIF4E levels are elevated in all ductal carcinoma *in situ* specimens and invasive ductal carcinomas, compared with benign breast specimens evaluated with Western blot analysis (64, 65). Preliminary studies suggest that this overexpression is attributable to gene amplification (66).

There are accumulating data suggesting that eIF4E overexpression can be valuable as a prognostic marker. eIF4E overexpression was found in a retrospective study to be a marker of poor prognosis in stages I to III breast carcinoma (67). Verification of the prognostic value of eIF4E in breast cancer is now under way in a prospective trial (67). However, in a different study, eIF4E expression was correlated with the aggressive behavior of non-Hodgkin's lymphomas (68). In a prospective analysis of patients with head and neck cancer, elevated levels of eIF4E in histologically tumor-free surgical margins predicted a significantly increased risk of local-regional recurrence (9). These results all suggest that eIF4E overexpression can be used to select patients who might benefit from more aggressive systemic therapy. Furthermore, the head and neck cancer data suggest that eIF4E overexpression is a field defect and can be used to guide local therapy.

Alterations in Other Initiation Factors. Alterations in a number of other initiation factors have been associated with cancer. Overproduction of eIF4G, similar to eIF4E, leads to malignant transformation *in vitro* (69). eIF-2α is found in increased levels in bronchioloalveolar carcinomas of the lung (3). Initiation factor eIF-4A1 is overexpressed in melanoma (70) and hepatocellular carcinoma (71). The p40 subunit of translation initiation factor 3 is amplified and overexpressed in breast and prostate cancer (72), and the eIF3-p110 subunit is overexpressed in testicular seminoma (73). The role that overexpression of these initiation factors plays on the development and progression of cancer, if any, is not known.

Overexpression of S6K. S6K is amplified and highly overexpressed in the MCF7 breast cancer cell line, compared with normal mammary epithelium (74). In a study by Barlund *et al.* (74), S6K was amplified in 59 of 668 primary breast tumors, and a statistically significant association was observed between amplification and poor prognosis.

Overexpression of PAP. PAP catalyzes 3′ poly(A) synthesis. PAP is overexpressed in human cancer cells compared with normal and virally transformed cells (75). PAP enzymatic activity in breast tumors has been correlated with PAP protein levels (76) and, in mammary tumor cytosols, was found to be an independent factor for predicting survival (76). Little is known, however, about how PAP expression or activity affects the translational profile.

Alterations in RNA-binding Proteins. Even less is known about alterations in RNA packaging in cancer. Increased expression and nuclear localization of the RNA-binding protein YB-1 are indicators of a poor prognosis for breast cancer (77), non-small cell lung cancer (78), and ovarian cancer (79). However, this effect may be mediated at least in part at the level of transcription, because YB-1 increases chemoresistance by enhancing the transcription of a multidrug resistance gene (80).

Activation of Signal Transduction Pathways

Activation of signal transduction pathways by loss of tumor suppressor genes or overexpression of certain tyrosine kinases can contribute to the growth and aggressiveness of tumors. An important mutant in human cancers is the tumor suppressor gene *PTEN*, which leads to the activation of the PI3K/Akt pathway. Activation of PI3K and Akt induces the oncogenic transformation of chicken embryo fibroblasts. The transformed cells show constitutive phosphorylation of S6K and of 4E-BP1 (81). A mutant Akt that retains kinase activity but does not phosphorylate S6K or 4E-BP1 does not transform fibroblasts, which suggests a correlation between the oncogenicity of PI3K and Akt and the phosphorylation of S6K and 4E-BP1 (81).

Several tyrosine kinases such as platelet-derived growth factor, insulin-like growth factor, HER2/neu, and epidermal growth factor receptor are overexpressed in cancer. Because these kinases activate downstream signal transduction pathways known to alter translation initiation, activation of translation is likely to contribute to the growth and aggressiveness of these tumors. Furthermore, the mRNA for many of these kinases themselves are under translational control. For example, HER2/neu mRNA is translationally controlled both by a short upstream open reading frame that represses HER2/neu translation in a cell type-independent manner and by a distinct cell type-dependent mechanism that increases translational efficiency (82). HER2/neu translation is different in transformed and normal cells. Thus, it is possible that alterations at the translational level can in part account for the discrepancy between *HER2/neu* gene amplification detected by fluorescence *in situ* hybridization and protein levels detected by immunohistochemical assays.

Translation Targets of Selected Cancer Therapy

Components of the translation machinery and signal pathways involved in the activation of translation initiation represent good targets for cancer therapy.

Targeting the mTOR Signaling Pathway: Rapamycin and Tumstatin

Rapamycin inhibits the proliferation of lymphocytes. It was initially developed as an immunosuppressive drug for organ transplantation. Rapamycin with FKBP 12 (FK506-binding protein, M_r 12,000) binds to mTOR to inhibit its function.

Rapamycin causes a small but significant reduction in the initiation rate of protein synthesis (83). It blocks cell growth in part by blocking S6 phosphorylation and selectively suppressing the translation of 5′ TOP mRNAs, such as ribosomal proteins, and elongation factors (83–85). Rapamycin also blocks 4E-BP1 phosphorylation and inhibits cap-dependent but not cap-independent translation (17, 86).

The rapamycin-sensitive signal transduction pathway, activated during malignant transformation and cancer progression, is now being studied as a target for cancer therapy (87). Prostate, breast, small cell lung, glioblastoma, melanoma, and T-cell leukemia are among the cancer lines most sensitive to the rapamycin analogue CCI-779 (Wyeth-Ayerst Research; Ref. 87). In rhabdomyosarcoma cell lines, rapamycin is either cytostatic or cytocidal, depending on the p53 status of the cell; p53 wild-type cells treated with rapamycin arrest in the G_1 phase and maintain their viability, whereas p53 mutant cells accumulate in G_1 and undergo apoptosis (88, 89). In a recently reported study using human primitive neuroectodermal tumor and medulloblastoma models, rapamycin exhibited more cytotoxicity in combination with cisplatin and camptothecin than as a single agent. *In vivo*, CCI-779 delayed growth of xenografts by 160% after 1 week of therapy and 240% after 2 weeks. A single high-dose administration caused a 37% decrease in tumor volume. Growth inhibition *in vivo* was 1.3 times greater, with cisplatin in combination with CCI-779 than with cisplatin alone (90). Thus, preclinical studies suggest that rapamycin analogues are useful as single agents and in combination with chemotherapy.

Rapamycin analogues CCI-779 and RAD001 (Novartis, Basel, Switzerland) are now in clinical trials. Because of the known effect of rapamycin on lymphocyte proliferation, a potential problem with rapamycin analogues is immunosuppression. However, although prolonged immunosuppression can result from rapamycin and CCI-779 administered on continuous-dose schedules, the immunosuppressive effects of rapamycin analogues resolve in ~24 h after therapy (91). The principal toxicities of CCI-779 have included dermatological toxicity, myelosuppression, infection, mucositis, diarrhea, reversible elevations in liver function tests, hyperglycemia, hypokalemia, hypocalcemia, and depression (87, 92–94). Phase II trials of CCI-779 have been conducted in advanced renal cell carcinoma and in stage III/IV breast carcinoma patients who failed with prior chemotherapy. In the results reported in abstract form, although there were no complete responses, partial responses were documented in both renal cell carcinoma and in breast carcinoma (94, 95). Thus, CCI-779 has documented preliminary clinical activity in a previously treated, unselected patient population.

Active investigation is under way into patient selection for mTOR inhibitors. Several studies have found an enhanced efficacy of CCI-779 in PTEN-null tumors (30, 96). Another study found that six of eight breast cancer cell lines were responsive to CCI-779, although only two of these lines lacked PTEN (97) There was, however, a positive correlation between Akt activation and CCI-779 sensitivity (97). This correlation suggests that activation of the PI3K-Akt pathway,

regardless of whether it is attributable to a PTEN mutation or to overexpression of receptor tyrosine kinases, makes cancer cell amenable to mTOR-directed therapy. In contrast, lower levels of the target of mTOR, 4E-BP1, are associated with rapamycin resistance; thus, a lower 4E-BP1/eIF4E ratio may predict rapamycin resistance (98).

Another mode of activity for rapamycin and its analogues appears to be through inhibition of angiogenesis. This activity may be both through direct inhibition of endothelial cell proliferation as a result of mTOR inhibition in these cells or by inhibition of translation of such proangiogenic factors as vascular endothelial growth factor in tumor cells (99, 100).

The angiogenesis inhibitor tumstatin, another anticancer drug currently under study, was also found recently to inhibit translation in endothelial cells (101). Through a requisite interaction with integrin, tumstatin inhibits activation of the PI3K/Akt pathway and mTOR in endothelial cells and prevents dissociation of eIF4E from 4E-BP1, thereby inhibiting cap-dependent translation. These findings suggest that endothelial cells are especially sensitive to therapies targeting the mTOR-signaling pathway.

Targeting eIF2α: EPA, Clotrimazole, mda-7, and Flavonoids

EPA is an n-3 polyunsaturated fatty acid found in the fish-based diets of populations having a low incidence of cancer (102). EPA inhibits the proliferation of cancer cells (103), as well as in animal models (104, 105). It blocks cell division by inhibiting translation initiation (105). EPA releases Ca^{2+} from intracellular stores while inhibiting their refilling, thereby activating PKR. PKR, in turn phosphorylates and inhibits eIF2α, resulting in the inhibition of protein synthesis at the level of translation initiation. Similarly, clotrimazole, a potent antiproliferative agent *in vitro* and *in vivo*, inhibits cell growth through depletion of Ca^{2+} stores, activation of PKR, and phosphorylation of eIF2α (106). Consequently, clotrimazole preferentially decreases the expression of cyclins A, E, and D1, resulting in blockage of the cell cycle in G_1.

mda-7 is a novel tumor suppressor gene being developed as a gene therapy agent. Adenoviral transfer of *mda-7* (Ad-mda7) induces apoptosis in many cancer cells including breast, colorectal, and lung cancer (107–109). Ad-mda7 also induces and activates PKR, which leads to phosphorylation of eIF2α and induction of apoptosis (110).

Flavonoids such as genistein and quercetin suppress tumor cell growth. All three mammalian eIF2α kinases, PKR, heme-regulated inhibitor, and PERK/PEK, are activated by flavonoids, with phosphorylation of eIF2α and inhibition of protein synthesis (111).

Targeting eIF4A and eIF4E: Antisense RNA and Peptides

Antisense expression of eIF4A decreases the proliferation rate of melanoma cells (112). Sequestration of eIF4E by overexpression of 4E-BP1 is proapoptotic and decreases tumorigenicity (113, 114). Reduction of eIF4E with antisense RNA decreases soft agar growth, increases tumor latency, and increases the rates of tumor doubling times (7). Antisense eIF4E RNA treatment also reduces the expression of angiogenic factors (115) and has been proposed as a potential adjuvant therapy for head and neck cancers, particularly when elevated eIF4E is found in surgical margins. Small molecule inhibitors that bind the eIF4G/4E-BP1-binding domain of eIF4E are proapoptotic (116) and are also being actively pursued.

Exploiting Selective Translation for Gene Therapy

A different therapeutic approach that takes advantage of the enhanced cap-dependent translation in cancer cells is the use of gene therapy vectors encoding suicide genes with highly structured 5′ UTR. These mRNA would thus be at a competitive disadvantage in normal cells and not translate well, whereas in cancer cells, they would translate more efficiently. For example, the introduction of the 5′ UTR of fibroblast growth factor-2 5′ to the coding sequence of *herpes simplex virus type-1 thymidine kinase* gene, allows for selective translation of *herpes simplex virus type-1 thymidine kinase* gene in breast cancer cell lines compared with normal mammary cell lines and results in selective sensitivity to ganciclovir (117).

Toward the Future

Translation is a crucial process in every cell. However, several alterations in translational control occur in cancer. Cancer cells appear to need an aberrantly activated translational state for survival, thus allowing the targeting of translation initiation with surprisingly low toxicity. Components of the translational machinery, such as eIF4E, and signal transduction pathways involved in translation initiation, such mTOR, represent promising targets for cancer therapy. Inhibitors of the mTOR have already shown some preliminary activity in clinical trials. It is possible that with the development of better predictive markers and better patient selection, response rates to single-agent therapy can be improved. Similar to other cytostatic agents, however, mTOR inhibitors are most likely to achieve clinical utility in combination therapy. In the interim, our increasing understanding of translation initiation and signal transduction pathways promise to lead to the identification of new therapeutic targets in the near future.

Acknowledgments

We thank Gayle Nesom from The University of Texas M. D. Anderson Cancer Center Department of Scientific Publications for editorial assistance and Dr. Elmer Bernstam for assistance with manuscript preparation.

References

1. Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I., and Hellen, C. U. Molecular mechanisms of translation initiation in eukaryotes. Proc. Natl. Acad. Sci. USA, 98: 7029–7036, 2001.

2. Rosenwald, I. B., Kaspar, R., Rousseau, D., Gehrke, L., Leboulch, P., Chen, J. J., Schmidt, E. V., Sonenberg, N., and London, I. M. Eukaryotic translation initiation factor 4E regulates expression of cyclin D1 at transcriptional and post-transcriptional levels. J. Biol. Chem., 270: 21176–21180, 1995.

3. Rosenwald, I. B., Hutzler, M. J., Wang, S., Savas, L., and Fraire, A. E. Expression of eukaryotic translation initiation factors 4E and 2α is increased frequently in bronchioloalveolar but not in squamous cell carcinomas of the lung. Cancer (Phila.), 92: 2164–2171, 2001.

4. Darveau, A., Pelletier, J., and Sonenberg, N. Differential efficiencies of *in vitro* translation of mouse c-myc transcripts differing in the 5' untranslated region. Proc. Natl. Acad. Sci. USA, *82:* 2315–2319, 1985.

5. Kozak, M. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA, *83:* 2850–2854, 1986.

6. Koromilas, A. E., Lazaris-Karatzas, A., and Sonenberg, N. mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. EMBO J., *11:* 4153–4158, 1992.

7. Rinker-Schaeffer, C. W., Graff, J. R., De Benedetti, A., Zimmer, S. G., and Rhoads, R. E. Decreasing the level of translation initiation factor 4E with antisense RNA causes reversal of ras-mediated transformation and tumorigenesis of cloned rat embryo fibroblasts. Int. J. Cancer, *55:* 841–847, 1993.

8. Kevil, C. G., De Benedetti, A., Payne, D. K., Coe, L. L., Laroux, F. S., and Alexander, J. S. Translational regulation of vascular permeability factor by eukaryotic initiation factor 4E: implications for tumor angiogenesis. Int. J. Cancer, *65:* 785–790, 1996.

9. Nathan, C. A., Franklin, S., Abreo, F. W., Nassar, R., De Benedetti, A., and Glass, J. Analysis of surgical margins with the molecular marker eIF4E: a prognostic factor in patients with head and neck cancer. J. Clin. Oncol., *17:* 2909–2914, 1999.

10. Fukunaga, R., and Hunter, T. MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. EMBO J., *16:* 1921–1933, 1997.

11. Waskiewicz, A. J., Flynn, A., Proud, C. G., and Cooper, J. A. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. EMBO J., *16:* 1909–1920, 1997.

12. Wang, X., Flynn, A., Waskiewicz, A. J., Webb, B. L., Vries, R. G., Baines, I. A., Cooper, J. A., and Proud, C. G. The phosphorylation of eukaryotic initiation factor eIF4E in response to phorbol esters, cell stresses, and cytokines is mediated by distinct MAP kinase pathways. J. Biol. Chem., *273:* 9373–9377, 1998.

13. Pyronnet, S., Imataka, H., Gingras, A. C., Fukunaga, R., Hunter, T., and Sonenberg, N. Human eukaryotic translation initiation factor 4G (eIF4G) recruits Mnk1 to phosphorylate eIF4E. EMBO J., *18:* 270–279, 1999.

14. Kleijn, M., Scheper, G. C., Voorma, H. O., and Thomas, A. A. Regulation of translation initiation factors by signal transduction. Eur. J. Biochem., *253:* 531–544, 1998.

15. Raught, B., and Gingras, A. C. eIF4E activity is regulated at multiple levels. Int. J. Biochem. Cell Biol., *31:* 43–57, 1999.

16. Takeuchi, K., Shibamoto, S., Nagamine, K., Shigemori, I., Omura, S., Kitamura, N., and Ito, F. Signaling pathways leading to transcription and translation cooperatively regulate the transient increase in expression of c-Fos protein. J. Biol. Chem., *276:* 26077–26083, 2001.

17. Kawasome, H., Papst, P., Webb, S., Keller, G. M., Johnson, G. L., Gelfand, E. W., and Terada, N. Targeted disruption of p70(s6k) defines its role in protein synthesis and rapamycin sensitivity. Proc. Natl. Acad. Sci. USA, *95:* 5033–5038, 1998.

18. Christie, G. R., Hajduch, E., Hundal, H. S., Proud, C. G., and Taylor, P. M. Intracellular sensing of amino acids in *Xenopus laevis* oocytes stimulates p70 S6 kinase in a target of rapamycin-dependent manner. J. Biol. Chem., *277:* 9952–9957, 2002.

19. Hara, K., Yonezawa, K., Weng, Q. P., Kozlowski, M. T., Belham, C., and Avruch, J. Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism. J. Biol. Chem., *273:* 14484–14494, 1998.

20. Graves, L. M., Bornfeldt, K. E., Argast, G. M., Krebs, E. G., Kong, X., Lin, T. A., and Lawrence, J. C., Jr. cAMP- and rapamycin-sensitive regulation of the association of eukaryotic initiation factor 4E and the translational regulator PHAS-I in aortic smooth muscle cells. Proc. Natl. Acad. Sci. USA, *92:* 7222–7226, 1995.

21. Merrick, W. C., and Hershey, J. W. B. The pathway and mechanism of eukaryotic protein synthesis. *In:* J. W. B. Hershey and M. B. Mathews (eds.), Translational Control, pp. 31–69. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1996.

22. Kimball, S. R. Eukaryotic initiation factor eIF2. Int. J. Biochem. Cell Biol., *31:* 25–29, 1999.

23. Jagus, R., Joshi, B., and Barber, G. N. PKR, apoptosis and cancer. Int. J. Biochem. Cell Biol., *31:* 123–138, 1999.

24. Thomas, G., and Hall, M. N. TOR signalling and control of cell growth. Curr. Opin. Cell Biol., *9:* 782–787, 1997.

25. Gingras, A. C., Raught, B., and Sonenberg, N. Regulation of translation initiation by FRAP/mTOR. Genes Dev., *15:* 807–826, 2001.

26. Gingras, A. C., Gygi, S. P., Raught, B., Polakiewicz, R. D., Abraham, R. T., Hoekstra, M. F., Aebersold, R., and Sonenberg, N. Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. Genes Dev., *13:* 1422–1437, 1999.

27. Kumar, V., Pandey, P., Sabatini, D., Kumar, M., Majumder, P. K., Bharti, A., Carmichael, G., Kufe, D., and Kharbhanda, S. Functional interaction between RAFT1/FRAP/mTOR and protein kinase Cδ in the regulation of cap-dependent initiation of translation. EMBO J., *19:* 1087–1097, 2000.

28. Yang, D. Q., and Kastan, M. B. Participation of ATM in insulin signalling through phosphorylation of eIF-4E-binding protein 1. Nat. Cell. Biol., *2:* 893–898, 2000.

29. Liu, G., Zhang, Y., Bode, A. M., Ma, W. Y., and Dong, Z. Phosphorylation of 4E-BP1 is mediated by the p38/MSK1 pathway in response to UVB irradiation. J. Biol. Chem., *277:* 8810–8816, 2002.

30. Neshat, M. S., Mellinghoff, I. K., Tran, C., Stiles, B., Thomas, G., Petersen, R., Frost, P., Gibbons, J. J., Wu, H., and Sawyers, C. L. Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR. Proc. Natl. Acad. Sci. USA, *98:* 10314–10319, 2001.

31. Sekulic, A., Hudson, C. C., Homme, J. L., Yin, P., Otterness, D. M., Karnitz, L. M., and Abraham, R. T. A direct linkage between the phosphoinositide 3-kinase-AKT signaling pathway and the mammalian target of rapamycin in mitogen-stimulated and transformed cells. Cancer Res., *60:* 3504–3513, 2000.

32. Scott, P. H., and Lawrence, J. C., Jr. Attenuation of mammalian target of rapamycin activity by increased cAMP in 3T3-L1 adipocytes. J. Biol. Chem., *273:* 34496–34501, 1998.

33. Reynolds, T. T., Bodine, S. C., and Lawrence, J. C., Jr. Control of Ser2448 phosphorylation in the mammalian target of rapamycin by insulin and skeletal muscle load. J. Biol. Chem., *277:* 17657–17662, 2002.

34. Peterson, R. T., Beal, P. A., Comb, M. J., and Schreiber, S. L. FKBP12-rapamycin-associated protein (FRAP) autophosphorylates at serine 2481 under translationally repressive conditions. J. Biol. Chem., *275:* 7416–7423, 2000.

35. Peterson, R. T., Desai, B. N., Hardwick, J. S., and Schreiber, S. L. Protein phosphatase 2A interacts with the 70-kDa S6 kinase and is activated by inhibition of FKBP12-rapamycin-associated protein. Proc. Natl. Acad. Sci. USA, *96:* 4438–4442, 1999.

36. McGrew, L. L., Dworkin-Rastl, E., Dworkin, M. B., and Richter, J. D. Poly(A) elongation during *Xenopus* oocyte maturation is required for translational recruitment and is mediated by a short sequence element. Genes Dev., *3:* 803–815, 1989.

37. Sheets, M. D., Wu, M., and Wickens, M. Polyadenylation of c-mos mRNA as a control point in *Xenopus* meiotic maturation. Nature (Lond.), *374:* 511–516, 1995.

38. Varnum, S. M., and Wormington, W. M. Deadenylation of maternal mRNAs during *Xenopus* oocyte maturation does not require specific *cis*-sequences: a default mechanism for translational control. Genes Dev., *4:* 2278–2286, 1990.

39. Gallie, D. R. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev., *5:* 2108–2116, 1991.

40. Sachs, A. B., and Varani, G. Eukaryotic translation initiation: there are (at least) two sides to every story. Nat. Struct. Biol., *7:* 356–361, 2000.

41. Wolffe, A. P., and Meric, F. Coupling transcription to translation: a novel site for the regulation of eukaryotic gene expression. Int. J. Biochem. Cell Biol., *28:* 247–257, 1996.

42. Evdokimova, V. M., Wei, C. L., Sitikov, A. S., Simonenko, P. N., Lazarev, O. A., Vasilenko, K. S., Ustinov, V. A., Hershey, J. W., and Ovchinnikov, L. P. The major protein of messenger ribonucleoprotein particles in somatic cells is a member of the Y-box binding transcription factor family. J. Biol. Chem., *270:* 3186–3192, 1995.

43. Matsumoto, K., Meric, F., and Wolffe, A. P. Translational repression dependent on the interaction of the *Xenopus* Y-box protein FRGY2 with mRNA. Role of the cold shock domain, tail domain, and selective RNA sequence recognition. J. Biol. Chem., *271:* 22706–22712, 1996.

44. Evdokimova, V., Ruzanov, P., Imataka, H., Raught, B., Svitkin, Y., Ovchinnikov, L. P., and Sonenberg, N. The major mRNA-associated protein YB-1 is a potent 5′ cap-dependent mRNA stabilizer. EMBO J., *20:* 5491–5502, 2001.

45. Saito, H., Hayday, A. C., Wiman, K., Hayward, W. S., and Tonegawa, S. Activation of the c-*myc* gene by translocation: a model for translational control. Proc. Natl. Acad. Sci. USA, *80:* 7476–7480, 1983.

46. Nanbru, C., Lafon, I., Audigier, S., Gensac, M. C., Vagner, S., Huez, G., and Prats, A. C. Alternative translation of the proto-oncogene c-*myc* by an internal ribosome entry site. J. Biol. Chem., *272:* 32061–32066, 1997.

47. Stoneley, M., Paulin, F. E., Le Quesne, J. P., Chappell, S. A., and Willis, A. E. c-Myc 5′ untranslated region contains an internal ribosome entry segment. Oncogene, *16:* 423–428, 1998.

48. Paulin, F. E., West, M. J., Sullivan, N. F., Whitney, R. L., Lyne, L., and Willis, A. E. Aberrant translational control of the c-*myc* gene in multiple myeloma. Oncogene, *13:* 505–513, 1996.

49. Chappell, S. A., LeQuesne, J. P., Paulin, F. E., de Schoolmeester, M. L., Stoneley, M., Soutar, R. L., Ralston, S. H., Helfrich, M. H., and Willis, A. E. A mutation in the c-myc-IRES leads to enhanced internal ribosome entry in multiple myeloma: a novel mechanism of oncogene de-regulation. Oncogene, *19:* 4437–4440, 2000.

50. Signori, E., Bagni, C., Papa, S., Primerano, B., Rinaldi, M., Amaldi, F., and Fazio, V. M. A somatic mutation in the 5′UTR of *BRCA1* gene in sporadic breast cancer causes down-modulation of translation efficiency. Oncogene, *20:* 4596–4600, 2001.

51. Liu, L., Dilworth, D., Gao, L., Monzon, J., Summers, A., Lassam, N., and Hogg, D. Mutation of the CDKN2A 5′ UTR creates an aberrant initiation codon and predisposes to melanoma. Nat. Genet., *21:* 128–132, 1999.

52. Savitsky, K., Platzer, M., Uziel, T., Gilad, S., Sartiel, A., Rosenthal, A., Elroy-Stein, O., Shiloh, Y., and Rotman, G. Ataxia-telangiectasia: structural diversity of untranslated sequences suggests complex post-transcriptional regulation of *ATM* gene expression. Nucleic Acids Res., *25:* 1678–1684, 1997.

53. Brown, C. Y., Mize, G. J., Pineda, M., George, D. L., and Morris, D. R. Role of two upstream open reading frames in the translational control of oncogene *mdm2*. Oncogene, *18:* 5631–5637, 1999.

54. Sobczak, K., and Krzyzosiak, W. J. Structural determinants of BRCA1 translational regulation. J. Biol. Chem., *277:* 17349–17358, 2002.

55. Arrick, B. A., Grendell, R. L., and Griffin, L. A. Enhanced translational efficiency of a novel transforming growth factor β3 mRNA in human breast cancer cells. Mol. Cell. Biol., *14:* 619–628, 1994.

56. Myers, S. M., Eng, C., Ponder, B. A., and Mulligan, L. M. Characterization of *RET* proto-oncogene 3′ splicing variants and polyadenylation sites: a novel C-terminus for RET. Oncogene, *11:* 2039–2045, 1995.

57. Byrne, J. A., Tomasetto, C., Rouyer, N., Bellocq, J. P., Rio, M. C., and Basset, P. The tissue inhibitor of *metalloproteinases-3* gene in breast carcinoma: identification of multiple polyadenylation sites and a stromal pattern of expression. Mol. Med., *1:* 418–427, 1995.

58. Moscow, J. A., He, R., Gudas, J. M., and Cowan, K. H. Utilization of multiple polyadenylation signals in the human *RHOA* protooncogene. Gene (Amst.), *144:* 229–238, 1994.

59. Senterre-Lesenfants, S., Alag, A. S., and Sobel, M. E. Multiple mRNA species are generated by alternate polyadenylation from the human *calmodulin-1* gene. J. Cell. Biochem., *58:* 445–454, 1995.

60. Lazaris-Karatzas, A., Montine, K. S., and Sonenberg, N. Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5′ cap. Nature (Lond.), *345:* 544–547, 1990.

61. De Benedetti, A., and Rhoads, R. E. Overexpression of eukaryotic protein synthesis initiation factor 4E in HeLa cells results in aberrant growth and morphology. Proc. Natl. Acad. Sci. USA, *87:* 8212–8216, 1990.

62. Polunovsky, V. A., Rosenwald, I. B., Tan, A. T., White, J., Chiang, L., Sonenberg, N., and Bitterman, P. B. Translational control of programmed cell death: eukaryotic translation initiation factor 4E blocks apoptosis in growth-factor-restricted fibroblasts with physiologically expressed or deregulated Myc. Mol. Cell. Biol., *16:* 6573–6581, 1996.

63. Miyagi, Y., Sugiyama, A., Asai, A., Okazaki, T., Kuchino, Y., and Kerr, S. J. Elevated levels of eukaryotic translation initiation factor eIF-4E, mRNA in a broad spectrum of transformed cell lines. Cancer Lett., *91:* 247–252, 1995.

64. Kerekatte, V., Smiley, K., Hu, B., Smith, A., Gelder, F., and De Benedetti, A. The proto-oncogene/translation factor *eIF4E*: a survey of its expression in breast carcinomas. Int. J. Cancer, *64:* 27–31, 1995.

65. Li, B. D., Liu, L., Dawson, M., and De Benedetti, A. Overexpression of eukaryotic initiation factor 4E (eIF4E) in breast carcinoma. Cancer (Phila.), *79:* 2385–2390, 1997.

66. Sorrells, D. L., Black, D. R., Meschonat, C., Rhoads, R., De Benedetti, A., Gao, M., Williams, B. J., and Li, B. D. Detection of *eIF4E* gene amplification in breast cancer by competitive PCR. Ann. Surg. Oncol., *5:* 232–237, 1998.

67. Li, B. D., McDonald, J. C., Nassar, R., and De Benedetti, A. Clinical outcome in stage I to III breast carcinoma and eIF4E overexpression. Ann. Surg., *227:* 756–761; discussion, 761–763, 1998.

68. Wang, S., Rosenwald, I. B., Hutzler, M. J., Pihan, G. A., Savas, L., Chen, J. J., and Woda, B. A. Expression of the eukaryotic translation initiation factors 4E and 2α in non-Hodgkin's lymphomas. Am. J. Pathol., *155:* 247–255, 1999.

69. Fukuchi-Shimogori, T., Ishii, I., Kashiwagi, K., Mashiba, H., Ekimoto, H., and Igarashi, K. Malignant transformation by overproduction of translation initiation factor eIF4G. Cancer Res., *57:* 5041–5044, 1997.

70. Eberle, J., Krasagakis, K., and Orfanos, C. E. Translation initiation factor eIF-4A1 mRNA is consistently overexpressed in human melanoma cells *in vitro*. Int. J. Cancer, *71:* 396–401, 1997.

71. Shuda, M., Kondoh, N., Tanaka, K., Ryo, A., Wakatsuki, T., Hada, A., Goseki, N., Igari, T., Hatsuse, K., Aihara, T., Horiuchi, S., Shichita, M., Yamamoto, N., and Yamamoto, M. Enhanced expression of translation factor mRNAs in hepatocellular carcinoma. Anticancer Res., *20:* 2489–2494, 2000.

72. Nupponen, N. N., Porkka, K., Kakkola, L., Tanner, M., Persson, K., Borg, A., Isola, J., and Visakorpi, T. Amplification and overexpression of p40 subunit of eukaryotic translation initiation factor 3 in breast and prostate cancer. Am. J. Pathol., *154:* 1777–1783, 1999.

73. Rothe, M., Ko, Y., Albers, P., and Wernert, N. Eukaryotic initiation factor 3 p110 mRNA is overexpressed in testicular seminomas. Am. J. Pathol., *157:* 1597–1604, 2000.

74. Barlund, M., Forozan, F., Kononen, J., Bubendorf, L., Chen, Y., Bittner, M. L., Torhorst, J., Haas, P., Bucher, C., Sauter, G., Kallioniemi, O. P., and Kallioniemi, A. Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. J. Natl. Cancer Inst. (Bethesda), *92:* 1252–1259, 2000.

75. Topalian, S. L., Kaneko, S., Gonzales, M. I., Bond, G. L., Ward, Y., and Manley, J. L. Identification and functional characterization of neo-poly(A) polymerase, an RNA processing enzyme overexpressed in human tumors. Mol. Cell. Biol., *21:* 5614–5623, 2001.

76. Scorilas, A., Talieri, M., Ardavanis, A., Courtis, N., Dimitriadis, E., Yotis, J., Tsiapalis, C. M., and Trangas, T. Polyadenylate polymerase enzymatic activity in mammary tumor cytosols: a new independent prognostic marker in primary breast cancer. Cancer Res., *60:* 5427–5433, 2000.

77. Janz, M., Harbeck, N., Dettmar, P., Berger, U., Schmidt, A., Jurchott, K., Schmitt, M., and Royer, H. D. Y-box factor YB-1 predicts drug resistance and patient outcome in breast cancer independent of clinically relevant tumor biologic factors HER2, uPA and PAI-1. Int. J. Cancer, *97:* 278–282, 2002.

78. Shibahara, K., Sugio, K., Osaki, T., Uchiumi, T., Maehara, Y., Kohno, K., Yasumoto, K., Sugimachi, K., and Kuwano, M. Nuclear expression of the Y-box binding protein, YB-1, as a novel marker of disease progression in non-small cell lung cancer. Clin. Cancer Res., *7:* 3151–3155, 2001.

79. Kamura, T., Yahata, H., Amada, S., Ogawa, S., Sonoda, T., Kobayashi, H., Mitsumoto, M., Kohno, K., Kuwano, M., and Nakano, H. Is nuclear expression of Y box-binding protein-1 a new prognostic factor in ovarian serous adenocarcinoma? Cancer (Phila.), *85:* 2450–2454, 1999.

80. Bargou, R. C., Jurchott, K., Wagener, C., Bergmann, S., Metzner, S., Bommert, K., Mapara, M. Y., Winzer, K. J., Dietel, M., Dorken, B., and Royer, H. D. Nuclear localization and increased levels of transcription factor YB-1 in primary human breast cancers are associated with intrinsic *MDR1* gene expression. Nat. Med., *3:* 447–450, 1997.

81. Aoki, M., Blazek, E., and Vogt, P. K. A role of the kinase mTOR in cellular transformation induced by the oncoproteins P3k and Akt. Proc. Natl. Acad. Sci. USA, *98:* 136–141, 2001.

82. Child, S. J., Miller, M. K., and Geballe, A. P. Cell type-dependent and -independent control of HER-2/neu translation. Int. J. Biochem. Cell Biol., *31:* 201–213, 1999.

83. Jefferies, H. B., Reinhard, C., Kozma, S. C., and Thomas, G. Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. Proc. Natl. Acad. Sci. USA, *91:* 4441–4445, 1994.

84. Terada, N., Patel, H. R., Takase, K., Kohno, K., Naim, A. C., and Gelfand, E. W. Rapamycin selectively inhibits translation of mRNAs encoding elongation factors and ribosomal proteins. Proc. Natl. Acad. Sci. USA, *91:* 11477–11481, 1994.

85. Jefferies, H. B., Fumagalli, S., Dennis, P. B., Reinhard, C., Pearson, R. B., and Thomas, G. Rapamycin suppresses 5'TOP mRNA translation through inhibition of p70s6k. EMBO J., *16:* 3693–3704, 1997.

86. Beretta, L., Gingras, A. C., Svitkin, Y. V., Hall, M. N., and Sonenberg, N. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. EMBO J., *15:* 658–664, 1996.

87. Hidalgo, M., and Rowinsky, E. K. The rapamycin-sensitive signal transduction pathway as a target for cancer therapy. Oncogene, *19:* 6680–6686, 2000.

88. Hosoi, H., Dilling, M. B., Shikata, T., Liu, L. N., Shu, L., Ashmun, R. A., Germain, G. S., Abraham, R. T., and Houghton, P. J. Rapamycin causes poorly reversible inhibition of mTOR and induces p53-independent apoptosis in human rhabdomyosarcoma cells. Cancer Res., *59:* 886–894, 1999.

89. Huang, S., and Houghton, P. J. Resistance to rapamycin: a novel anticancer drug. Cancer Metastasis Rev., *20:* 69–78, 2001.

90. Geoerger, B., Kerr, K., Tang, C. B., Fung, K. M., Powell, B., Sutton, L. N., Phillips, P. C., and Janss, A. J. Antitumor activity of the rapamycin analog CCI-779 in human primitive neuroectodermal tumor/medulloblastoma models as single agent and in combination chemotherapy. Cancer Res., *61:* 1527–1532, 2001.

91. Gibbons, J. J., Discafani, C., Peterson, R., Hernandez, R., Skotnicki, J., and Frost, P. The effect of CCI-779, a novel macrolide anti-tumor agent, on the growth of human tumor cells *in vitro* and in nude mouse xenografts *in vivo*. Proc. Am. Assoc. Cancer Res., *40:* 301, 1999.

92. Hidalgo, M., Rowinsky, E., Erlichman, C., Marshall, B., Marks, R., Edwards, T., and Buckner, J. J. A Phase I and pharmacological study of CCI-779 cycle inhibitor. Ann. Oncol., *11* (Suppl. 4): 133, 2001.

93. Alexandre, J., Raymond, E., Depenbrock, H., Mekhaldi, S., Angevin, E., Paillet, C., Hanauske, A., Frisch, J., Feussner, A., and Armand, J. P. CCI-779, a new rapamycin analog, has antitumor activity at doses inducing only mild cutaneous effects and mucositis: early results of an ongoing Phase I study. Proceedings of the 1999 AACR-NCI-EORTC International Conference, Clin. Cancer Res., *5* (Suppl.): 3730s, 1999.

94. Chan, S., Johnston, S., Scheulen, M. E., Mross, K., Morant, A., Lahr, A., Feussner, A., Berger, M., and Kirsch, T. First report: a Phase 2 study of the safety and activity of CCI-779 for patients with locally advanced or metastatic breast cancer failing prior chemotherapy. Proc. Am. Soc. Clin. Oncol., *21:* 44a, 2002.

95. Atkins, M. B., Hidalgo, M., Stadler, W., Logan, T., Dutcher, J. P., Hudes, G., Park, Y., Marshall, B., Boni, J., and Dukart, G. A randomized double-blind Phase 2 study of intravenous CCI-779 administered weekly to patients with advanced renal cell carcinoma. Proc. Am. Soc. Clin. Oncol., *21:* 10a, 2002.

96. Smith, S. G., Trinh, C. M., Inge, L. J., Thomas, G., Cloughsey, T. F., Sawyers, C. L., and Mischel, P. S. PTEN expression status predicts glioblastoma cell sensitivity to CCI-779. Proc. Am. Assoc. Cancer Res., *43:* 335, 2002.

97. Yu, K., Toral-Barza, L., Discafani, C., Zhang, W. G., Skotnicki, J., Frost, P., and Gibbons, J. J. mTOR, a novel target in breast cancer: the effect of CCI-779, an mTOR inhibitor, in preclinical models of breast cancer. Endocr. Relat. Cancer, *8:* 249–258, 2001.

98. Dilling, M. B., Germain, G. S., Dudkin, L., Jayaraman, A. L., Zhang, X., Harwood, F. C., and Houghton, P. J. 4E-binding proteins, the suppressors of eukaryotic initiation factor 4E, are downregulated in cells with acquired or intrinsic resistance to rapamycin. J. Biol. Chem., *277:* 13907–13917, 2002.

99. Guba, M., von Breitenbuch, P., Steinbauer, M., Koehl, G., Flegel, S., Hornung, M., Bruns, C. J., Zuelke, C., Farkas, S., Anthuber, M., Jauch, K. W., and Geissler, E. K. Rapamycin inhibits primary and metastatic tumor growth by antiangiogenesis: involvement of vascular endothelial growth factor. Nat. Med., *8:* 128–135, 2002.

100. Lane, H. A., Schell, C., Theuer, A., O'Reilly, T., and Wood, J. Anti-angiogenic activity of RAD001, an orally active anticancer agent. Proc. Am. Assoc. Cancer Res., *43:* 184, 2002.

101. Masshima, Y., Sudhakar, A., Lively, J. C., Ueki, K., Kharbanda, S., Kahn, C. R., Sonenberg, N., Hynes, R. O., and Kalluri, R. Tumstatin, an endothelial cell-specific inhibitor of protein synthesis. Science (Wash. DC), *295:* 140–143, 2002.

102. Caygill, C. P., Charlett, A., and Hill, M. J. Fat, fish, fish oil and cancer. Br. J. Cancer, *74:* 159–164, 1996.

103. Falconer, J. S., Ross, J. A., Fearon, K. C., Hawkins, R. A., O'Riordain, M. G., and Carter, D. C. Effect of eicosapentaenoic acid and other fatty acids on the growth *in vitro* of human pancreatic cancer cell lines. Br. J. Cancer, *69:* 826–832, 1994.

104. Noguchi, M., Minami, M., Yagasaki, R., Kinoshita, K., Earashi, M., Kitagawa, H., Taniya, T., and Miyazaki, I. Chemoprevention of DMBA-induced mammary carcinogenesis in rats by low-dose EPA and DHA. Br. J. Cancer, *75:* 348–353, 1997.

105. Palakurthi, S. S., Fluckiger, R., Aktas, H., Changolkar, A. K., Shahsafaei, A., Harnett, S., Kilic, E., and Halperin, J. A. Inhibition of translation initiation mediates the anticancer effect of the n-3 polyunsaturated fatty acid eicosapentaenoic acid. Cancer Res., *60:* 2919–2925, 2000.

106. Aktas, H., Fluckiger, R., Acosta, J. A., Savage, J. M., Palakurthi, S. S., and Halperin, J. A. Depletion of intracellular Ca^{2+} stores, phosphorylation of eIF2α, and sustained inhibition of translation initiation mediate the anticancer effects of clotrimazole. Proc. Natl. Acad. Sci. USA, *95:* 8280–8285, 1998.

107. Mhashilkar, A. M., Schrock, R. D., Hindi, M., Liao, J., Sieger, K., Kourouma, F., Zou-Yang, X. H., Onishi, E., Takh, O., Vedvick, T. S., Fanger, G., Stewart, L., Watson, G. J., Snary, D., Fisher, P. B., Saeki, T., Roth, J. A., Ramesh, R., and Chada, S. Melanoma differentiation associated gene-7 (*mda-7*): a novel anti-tumor gene for cancer gene therapy. Mol. Med., *7:* 271–282, 2001.

108. Su, Z. Z., Madireddi, M. T., Lin, J. J., Young, C. S., Kitada, S., Reed, J. C., Goldstein, N. I., and Fisher, P. B. The cancer growth suppressor gene *mda-7* selectively induces apoptosis in human breast cancer cells and inhibits tumor growth in nude mice. Proc. Natl. Acad. Sci. USA, *95:* 14400–14405, 1998.

109. Saeki, T., Mhashilkar, A., Chada, S., Branch, C., Roth, J. A., and Ramesh, R. Tumor-suppressive effects by adenovirus-mediated *mda-7* gene transfer in non-small cell lung cancer cell *in vitro*. Gene Ther., *7:* 2051–2057, 2000.

110. Pataer, A., Vorburger, S. A., Barber, G. N., Chada, S., Mhashilkar, A. M., Zou-Yang, H., Stewart, A. L., Balachandran, S., Roth, J. A., Hunt, K. K., and Swisher, S. G. Adenoviral transfer of the melanoma differentiation-associated gene 7 (*mda7*) induces apoptosis of lung cancer cells via up-regulation of the double-stranded RNA-dependent protein kinase (PKR). Cancer Res., *62:* 2239–2243, 2002.

111. Ito, T., Warnken, S. P., and May, W. S. Protein synthesis inhibition by flavonoids: roles of eukaryotic initiation factor 2α kinases. Biochem. Biophys. Res. Commun., *265:* 589–594, 1999.

112. Eberle, J., Fecker, L. F., Bittner, J. U., Orfanos, C. E., and Geilen, C. C. Decreased proliferation of human melanoma cell lines caused by antisense RNA against translation factor eIF-4A1. Br. J. Cancer, *86:* 1957–1962, 2002.

113. Polunovsky, V. A., Gingras, A. C., Sonenberg, N., Peterson, M., Tan, A., Rubins, J. B., Manivel, J. C., and Bitterman, P. B. Translational control of the antiapoptotic function of Ras. J. Biol. Chem., *275:* 24776–24780, 2000.

114. D'Cunha, J., Kratzke, M. G., Alter, M. D., Polunovsky, V. A., Bitterman, P. B., and Kratzke, R. A. Over-expression of the translational repressor 4E-BP1 inhibits NSCLC tumorigenicity *in vivo*. Proc. Am. Assoc. Cancer Res., *43:* 816–817, 2002.

115. DeFatta, R. J., Nathan, C. A., and De Benedetti, A. Antisense RNA to eIF4E suppresses oncogenic properties of a head and neck squamous cell carcinoma cell line. Laryngoscope, *110:* 928–933, 2000.

116. Herbert, T. P., Fahraeus, R., Prescott, A., Lane, D. P., and Proud, C. G. Rapid induction of apoptosis mediated by peptides that bind initiation factor eIF4E. Curr. Biol., *10:* 793–796, 2000.

117. DeFatta, R. J., Li, Y., and De Benedetti, A. Selective killing of cancer cells based on translational control of a suicide gene. Cancer Gene Ther., *9:* 573–578, 2002.

EXHIBIT 9

MOLECULAR AND CELLULAR BIOLOGY, Mar. 1999, p. 1720–1730 Vol. 19, No. 3
0270-7306/99/$04.00+0

Correlation between Protein and mRNA Abundance in Yeast

STEVEN P. GYGI, YVAN ROCHON, B. ROBERT FRANZA, AND RUEDI AEBERSOLD*
Department of Molecular Biotechnology, University of Washington, Seattle, Washington 98195-7730

Received 5 October 1998/Returned for modification 11 November 1998/Accepted 2 December 1998

We have determined the relationship between mRNA and protein expression levels for selected genes expressed in the yeast *Saccharomyces cerevisiae* growing at mid-log phase. The proteins contained in total yeast cell lysate were separated by high-resolution two-dimensional (2D) gel electrophoresis. Over 150 protein spots were excised and identified by capillary liquid chromatography-tandem mass spectrometry (LC-MS/MS). Protein spots were quantified by metabolic labeling and scintillation counting. Corresponding mRNA levels were calculated from serial analysis of gene expression (SAGE) frequency tables (V. E. Velculescu, L. Zhang, W. Zhou, J. Vogelstein, M. A. Basrai, D. E. Bassett, Jr., P. Hieter, B. Vogelstein, and K. W. Kinzler, Cell 88:243–251, 1997). We found that the correlation between mRNA and protein levels was insufficient to predict protein expression levels from quantitative mRNA data. Indeed, for some genes, while the mRNA levels were of the same value the protein levels varied by more than 20-fold. Conversely, invariant steady-state levels of certain proteins were observed with respective mRNA transcript levels that varied by as much as 30-fold. Another interesting observation is that codon bias is not a predictor of either protein or mRNA levels. Our results clearly delineate the technical boundaries of current approaches for quantitative analysis of protein expression and reveal that simple deduction from mRNA transcript analysis is insufficient.

The description of the state of a biological system by the quantitative measurement of the system constituents is an essential but largely unexplored area of biology. With recent technical advances including the development of differential display-PCR (21), of cDNA microarray and DNA chip technology (20, 27), and of serial analysis of gene expression (SAGE) (34, 35), it is now feasible to establish global and quantitative mRNA expression profiles of cells and tissues in species for which the sequence of all the genes is known. However, there is emerging evidence which suggests that mRNA expression patterns are necessary but are by themselves insufficient for the quantitative description of biological systems. This evidence includes discoveries of posttranscriptional mechanisms controlling the protein translation rate (15), the half-lives of specific proteins or mRNAs (33), and the intracellular location and molecular association of the protein products of expressed genes (32).

Proteome analysis, defined as the analysis of the protein complement expressed by a genome (26), has been suggested as an approach to the quantitative description of the state of a biological system by the quantitative analysis of protein expression profiles (36). Proteome analysis is conceptually attractive because of its potential to determine properties of biological systems that are not apparent by DNA or mRNA sequence analysis alone. Such properties include the quantity of protein expression, the subcellular location, the state of modification, and the association with ligands, as well as the rate of change with time of such properties. In contrast to the genomes of a number of microorganisms (for a review, see reference 11) and the transcriptome of *Saccharomyces cerevisiae* (35), which have been entirely determined, no proteome map has been completed to date.

The most common implementation of proteome analysis is the combination of two-dimensional gel electrophoresis (2DE) (isoelectric focusing-sodium dodecyl sulfate [SDS]-polyacrylamide gel electrophoresis) for the separation and quantitation of proteins with analytical methods for their identification. 2DE permits the separation, visualization, and quantitation of thousands of proteins reproducibly on a single gel (18, 24). By itself, 2DE is strictly a descriptive technique. The combination of 2DE with protein analytical techniques has added the possibility of establishing the identities of separated proteins (1, 2) and thus, in combination with quantitative mRNA analysis, of correlating quantitative protein and mRNA expression measurements of selected genes.

The recent introduction of mass spectrometric protein analysis techniques has dramatically enhanced the throughput and sensitivity of protein identification to a level which now permits the large-scale analysis of proteins separated by 2DE. The techniques have reached a level of sensitivity that permits the identification of essentially any protein that is detectable in the gels by conventional protein staining (9, 29). Current protein analytical technology is based on the mass spectrometric generation of peptide fragment patterns that are idiotypic for the sequence of a protein. Protein identity is established by correlating such fragment patterns with sequence databases (10, 22, 37). Sophisticated computer software (8) has automated the entire process such that proteins are routinely identified with no human interpretation of peptide fragment patterns.

In this study, we have analyzed the mRNA and protein levels of a group of genes expressed in exponentially growing cells of the yeast *S. cerevisiae*. Protein expression levels were quantified by metabolic labeling of the yeast proteins to a steady state, followed by 2DE and liquid scintillation counting of the selected, separated protein species. Separated proteins were identified by in-gel tryptic digestion of spots with subsequent analysis by microspray liquid chromatography-tandem mass spectrometry (LC-MS/MS) and sequence database searching. The corresponding mRNA transcript levels were calculated from SAGE frequency tables (35).

This study, for the first time, explores a quantitative comparison of mRNA transcript and protein expression levels for a relatively large number of genes expressed in the same metabolic state. The resultant correlation is insufficient for predic-

* Corresponding author. Mailing address: Department of Molecular Biotechnology, Box 357730, University of Washington, Seattle, WA 98195-7730. Phone: (206) 221-4196. Fax: (206) 685-7301. E-mail: ruedi@u.washington.edu.

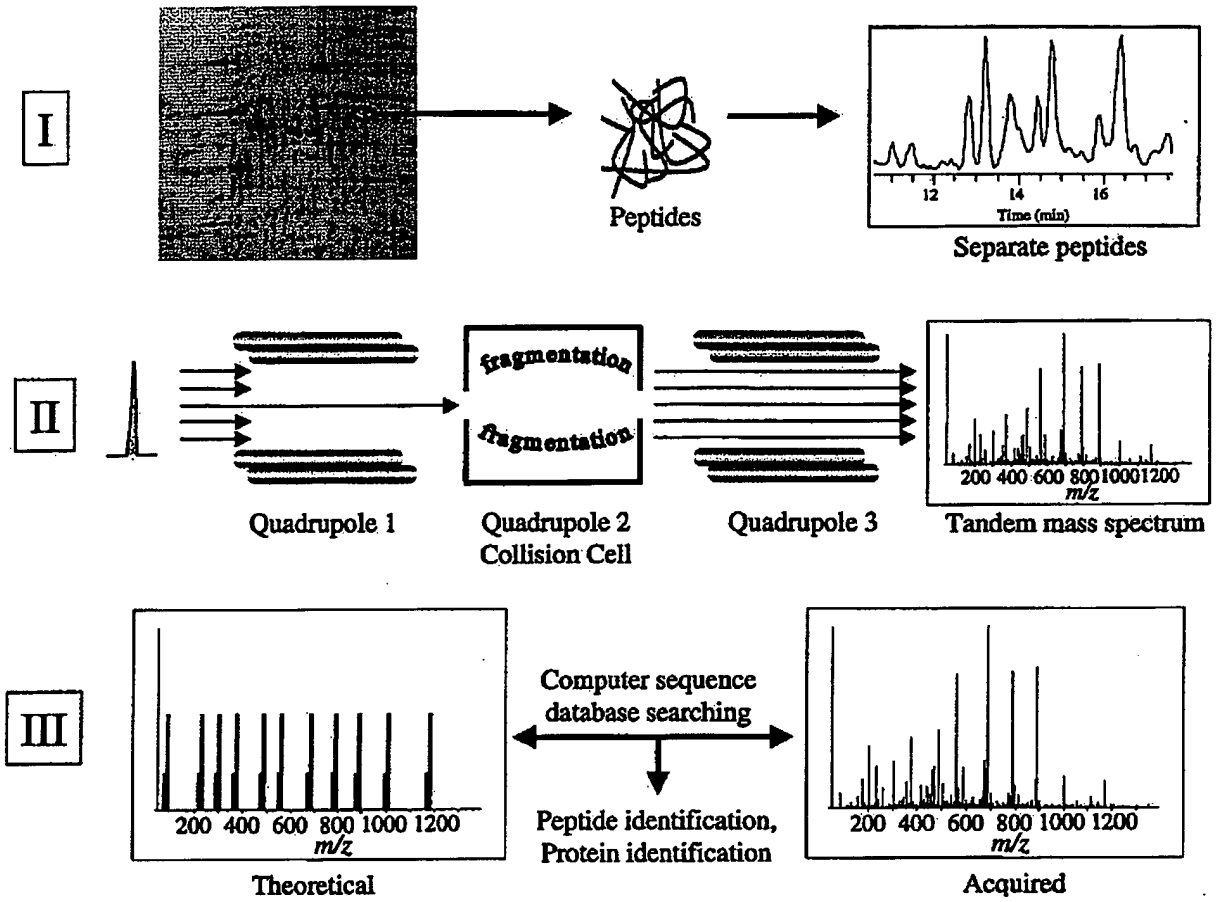


FIG. 1. Schematic illustration of proteome analysis by 2DE and mass spectrometry. In part I, proteins are separated by 2DE, stained spots are excised and subjected to in-gel digestion with trypsin, and the resulting peptides are separated by on-line capillary high-performance liquid chromatography. In part II, a peptide is shown eluting from the column in part I. The peptide is ionized by electrospray ionization and enters the mass spectrometer. The mass of the ionized peptide is detected, and the first quadrupole mass filter allows only the specific mass-to-charge ratio of the selected peptide ion to pass into the collision cell. In the collision cell, the energized, ionized peptides collide with neutral argon gas molecules. Fragmentation of the peptide is essentially random but occurs mainly at the peptide bonds, resulting in smaller peptides of differing lengths (masses). These peptide fragments are detected as a tandem mass (MS/MS) spectrum in the third quadrupole mass filter where two ion series are recorded simultaneously, one each from sequencing inward from the N and C termini of the peptide, respectively. In part III, the MS/MS spectrum from the selected, ionized peptide is compared to predicted tandem mass spectra computer generated from a sequence database. Provided that the peptide sequence exists in the database, the peptide and, by association, the protein from which the peptide was derived can be identified. Unambiguous protein identification is attained in a single analysis because multiple peptides are identified as being derived from the same protein.

tion of protein levels from mRNA transcript levels. We have also compared the relative amounts of protein and mRNA with the respective codon bias values for the corresponding genes. This comparison indicates that codon bias by itself is insufficient to accurately predict either the mRNA or the protein expression levels of a gene. In addition, the results demonstrate that only highly expressed proteins are detectable by 2DE separation of total cell lysates and that therefore the construction of complete proteome maps with current technology will be very challenging, irrespective of the type of organism.

MATERIALS AND METHODS

Yeast strain and growth conditions. The source of protein and message transcripts for all experiments was YPH499 (*MATa ura3-52 lys2-801 ade2-101 leu2-Δ1 his3-Δ200 trp1-Δ63*) (30). Logarithmically growing cells were obtained by growing yeast cells to early log phase (3×10^6 cells/ml) in YPD rich medium (YPD supplemented with 6 mM uracil, 4.8 mM adenine, and 24 mM tryptophan) at 30°C (35). Metabolic labeling of protein was accomplished in YPD medium exactly as described elsewhere (4) with the exception that 1 ml of cells was labeled with 3 mCi to offset methionine present in YPD medium. Protein was harvested as described by Garrels and coworkers (12). Harvested protein was lyophilized, resuspended in isoelectric focusing gel rehydration solution, and stored at −80°C.

2DE. Soluble proteins were run in the first dimension by using a commercial flatbed electrophoresis system (Multiphor II; Pharmacia Biotech). Immobilized polyacrylamide gel (IPG) dry strips with nonlinear pH 3.0 to 10.0 gradients (Amersham-Pharmacia Biotech) were used for the first-dimension separation. Forty micrograms of protein from whole-cell lysates was mixed with IPG strip rehydration buffer (8 M urea, 2% Nonidet P-40, 10 mM dithiothreitol), and 250 to 380 μl of solution was added to individual lanes of an IPG strip rehydration tray (Amersham-Pharmacia Biotech). The strips were allowed to rehydrate at room temperature for 1 h. The samples were run at 300 V–10 mA–5 W for 2 h, then ramped to 3,500 V–10 mA–5 W over a period of 3 h, and then kept at 3,500 V–10 mA–5 W for 15 to 19 h. At the end of the first-dimension run (60 to 70 kV · h), the IPG strips were recquilibrated for 8 min in 2% (wt/vol) dithiothreitol in 2% (wt/vol) SDS–6 M urea–30% (wt/vol) glycerol–0.05 M Tris HCl (pH 6.8) and for 4 min in 2.5% iodoacetamide in 2% (wt/vol) SDS–6 M urea–30% (wt/vol) glycerol–0.05 M Tris HCl (pH 6.8). Following recquilibration, the strips were transferred and apposed to 10% polyacrylamide second-dimension gels. Polyacrylamide gels were poured in a casting stand with 10% acrylamide–2.67% piperazine diacrylamide–0.375 M Tris base-HCl (pH 8.8)–0.1% (wt/vol) SDS–0.05%

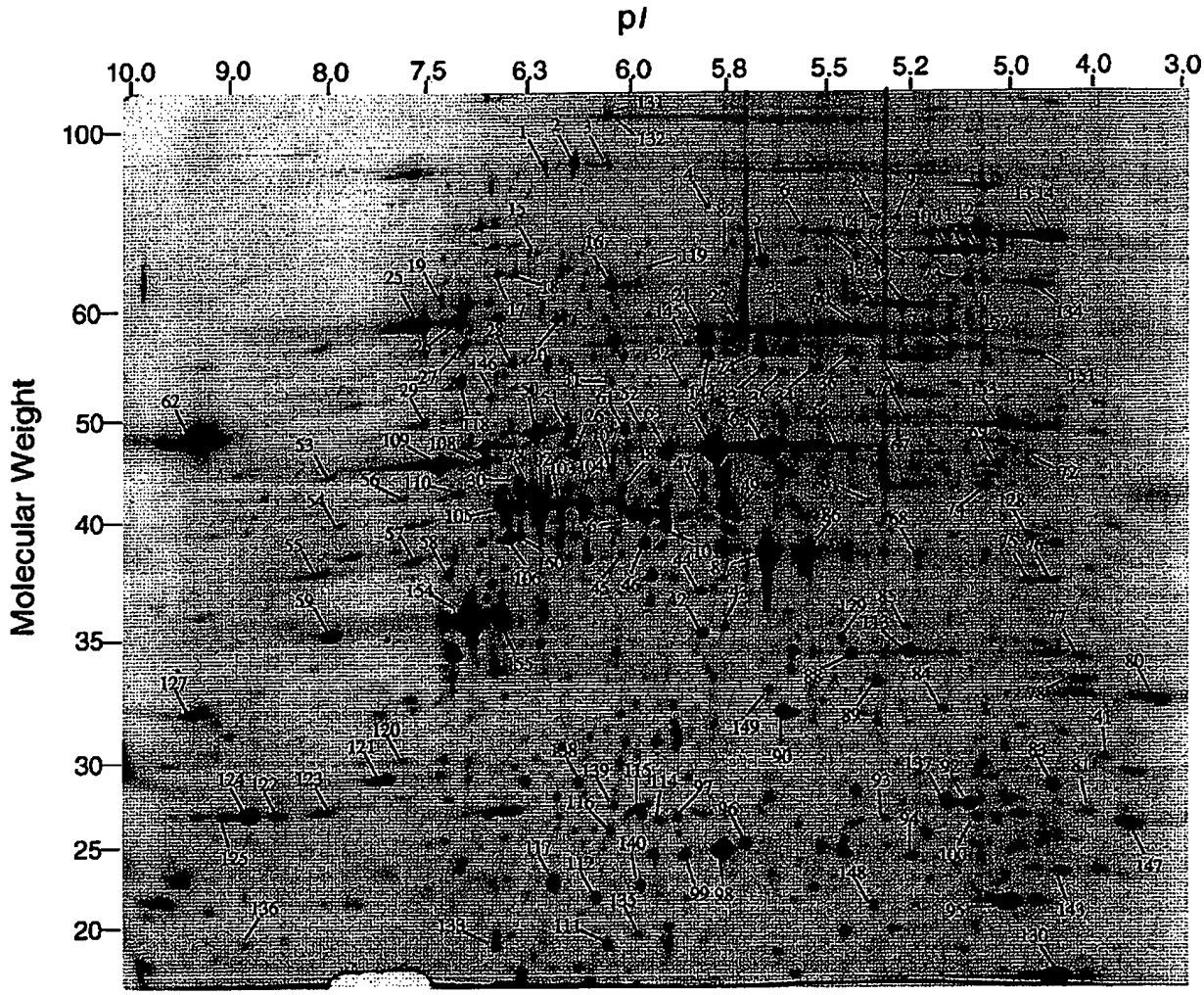


FIG. 2. 2D silver-stained gel of the proteins in yeast total cell lysate. Proteins were separated in the first dimension (horizontal) by isoelectric focusing and then in the second dimension (vertical) by molecular weight sieving. Protein spots (156) were chosen to include the entire range of molecular weights, isoelectric focusing points, and staining intensities. Spots were excised, and the corresponding protein was identified by mass spectrometry and database searching. The spots are labeled on the gel and correspond to the data presented in Table 1. Molecular weights are given in thousands.

(wt/vol) ammonium persulfate–0.05% TEMED (*N,N,N',N'*-tetramethylethylenediamine) in Milli-Q water. The apparatus used to run second-dimension gels was a noncommercial apparatus from Oxford Glycosciences, Inc. Once the IPG strips were apposed to the second-dimension gels, they were immediately run at 50 mA (constant)–500 V–85 W for 20 min, followed by 200 mA (constant)–500 V–85 W until the buffer front line was 10 to 15 mm from the bottom of the gel. Gels were removed and silver stained according to the procedure of Shevchenko et al. (29).

Protein identification. Gels were exposed to X-ray film overnight, and then the silver staining and film were used to excise 156 spots of varying intensities, molecular weights, and isoelectric focusing points. In order to increase the detection limit by mass spectrometry, spots were cut out and pooled from up to four identical cold, silver-stained gels. In-gel tryptic digests of pooled spots were performed as described previously (29). Tryptic peptides were analyzed by microcapillary LC-MS with automated switching to MS/MS mode for peptide fragmentation. Spectra were searched against the composite OWL protein sequence database (version 30.2; 250,514 protein sequences) (24a) by using the computer program Sequest (8), which matches theoretical and acquired tandem mass spectra. A protein match was determined by comparing the number of peptides identified and their respective cross-correlation scores. All protein identifications were verified by comparison with theoretical molecular weights and isoelectric points.

mRNA quantitation. Velculescu and coworkers have previously generated frequency tables for yeast mRNA transcripts from the same strain grown under the same stated conditions as described herein (35). The SAGE technology is based on two main principles. First, a short sequence tag (15 bp) that contains sufficient information uniquely to identify a transcript is generated. A single tag is usually generated from each mRNA transcript in the cell which corresponds to 15 bp at the 3'-most cutting site for *Nla*III. Second, many transcript tags can be concatenated into a single molecule and then sequenced, revealing the identity of multiple tags simultaneously. Over 20,000 transcripts were sequenced from yeast strain YPH499 growing at mid-log phase on glucose. Assuming the previously derived estimate of 15,000 mRNA molecules per cell (16), this would represent a 1.3-fold coverage even for mRNA molecules present at a single copy per cell and would provide a 72% probability of detecting such transcripts. Computer software which took for input the gene detected, examined the nucleotide sequence, and performed the calculation as described by Velculescu and coworkers (35) was written. In practice, we found that for 21 of 128 (16%) genes examined viable mRNA levels from SAGE data could not be calculated. This was because (i) no CATG site was found in the open reading frame (ORF), (ii) a CATG site was found but the corresponding 10-bp putative SAGE tag was not found in the frequency tables, or (iii) identical putative SAGE tags were present for multiple genes (e.g., TDH2_YEAST and TDH3_YEAST).

TABLE 1. Expressed genes identified from 2D gel in Fig. 2

Mol wt	pI	Spot no.	YPD gene name[a]	Protein abundance (10^3 copies/cell)	mRNA abundance (copies/cell)	Codon bias
17,259	6.75	133	CPR1	15.2	61.7	0.769
18,702	4.80	83	EGD2	20.1	5.2	0.724
18,726	4.44	147	YKL056C	61.2	88.4	0.831
18,978	5.95	135	YER067W	3.7	6.7	0.118
19,108	5.04	130	YLR109W	94.4	9.7	0.680
19,681	9.08	136	ATP7	11.0	NA[b,c]	0.246
20,505	6.07	111	GUK1	16.5	3.7	0.422
21,444	5.25	148	SAR1	5.4	10.4	0.455
21,583	4.98	95	TSA1	110.6	40.1	0.845
22,602	4.30	80	EFB1	66.1	23.8	0.875
23,079	6.29	112	SOD2	12.6	2.2	0.351
23,743	5.44	137	HSP26	NA[d]	0.7	0.434
24,033	5.97	96	ADK1	17.4	16.4	0.656
24,058	4.43	143	YKL117W	29.2	10.4	0.339
24,353	6.30	140	TFS1	8.1	0.7	0.146
24,662	5.85	99	URA5	25.4	6.0	0.359
24,808	6.33	97	GSP1	26.3	5.2	0.735
24,908	8.73	122	RPS5	18.6	NA[c]	0.899
25,081	4.65	81	MRP8	9.3	NA[c]	0.241
25,960	6.06	116	RPE1	5.8	0.7	0.372
26,378	9.55	127	RPS3	96.8	NA[c]	0.863
26,467	5.18	100	VMA4	10.5	3.7	0.427
26,661	5.84	98	TPI1	NA[d]	NA[c]	0.900
27,156	5.56	93	PRE8	6.9	0.7	0.129
27,334	6.13	115	YHR049W	18.4	2.2	0.520
27,472	5.33	92	YNL010W	31.6	3.7	0.421
27,480	8.95	123	GPM1	10.0	169.4	0.902
27,480	8.95	124	GPM1	231.4	169.4	0.902
27,480	8.95	125	GPM1	7.5	169.4	0.902
27,809	5.97	139	HOR2	5.7	0.7	0.381
27,874	4.46	78	YST1	13.6	52.8	0.805
28,595	4.51	41	PUP2	4.4	0.7	0.147
29,156	6.59	114	YMR226C	14.5	2.2	0.283
29,244	8.40	120	DPM1	5.0	11.2	0.362
29,443	5.91	48	PRE4	3.4	3.7	0.162
30,012	6.39	138	PRB1	21.2	1.5	0.449
30,073	4.63	77	BMH1	14.7	28.2	0.454
30,296	7.94	121	OMP2	67.4	41.6	0.499
30,435	6.34	89	GPP1	70.2	11.2	0.703
31,332	5.57	88	ILV6	13.9	3.0	0.402
32,159	5.46	113	IPP1	63.1	3.7	0.752
32,263	6.00	149	HIS1	22.4	4.5	0.232
33,311	5.35	84	SPE3	15.1	6.7	0.468
34,465	5.60	129	ADE1	8.7	5.2	0.305
34,762	5.32	85	SEC14	10.9	6.0	0.373
34,797	5.85	42	URA1	49.5	8.9	0.237
34,799	6.04	90	BEL1	103.2	81.0	0.875
35,556	5.97	43	YDL124W	6.4	4.5	0.206
35,619	8.41	59	TDH1	69.8	32.7[e]	0.940
35,650	5.49	68	CAR1	5.2	3.0	0.339
35,712	6.72	117	TDH2	49.6	473.0[e]	0.982
35,712	6.72	154	TDH2	863.5	473.0[e]	0.982
35,712	6.72	155	TDH2	79.4	473.0[e]	0.982
36,272	4.85	128	APA1	8.7	0.7	0.425
36,358	5.05	75	YJR105W	17.6	17.1	0.522
36,358	5.05	76	YJR105W	27.5	17.1	0.522
36,596	6.37	79	ADH2	58.9	260.0[e]	0.711
36,714	6.30	102	ADH1	746.1	260.0	0.913
36,714	6.30	103	ADH1	17.6	260.0	0.913
36,714	6.30	104	ADH1	61.4	260.0	0.913
36,714	6.30	105	ADH1	52.7	260.0	0.913
37,033	6.23	44	TAL1	44.8	3.7	0.701
37,796	7.36	57	IDH2	29.4	6.7	0.330
37,886	6.49	106	ILV5	76.0	4.5	0.892
38,700	7.83	55	BAT1	30.9	11.2	0.469
38,702	6.24	46	QCR2	NA[d]	2.2	0.326
39,477	5.58	86	FBA1	17.8	183.6	0.935
39,477	5.58	87	FBA1	427.2	183.6	0.935
39,540	6.50	150	HOM2	60.3	4.5	0.592
39,561	6.12	156	PSA1	96.4	27.5	0.718
41,158	6.01	49	YNL134C	14.9	1.5	0.316
41,623	7.18	58	BAT2	19.0	8.9	0.250
41,728	7.29	110	ERG10	24.1	4.5	0.543
41,900	5.42	74	TOM40	22.3	2.2	0.375
42,402	6.29	45	CYS3	6.7	8.9	0.621
42,883	5.63	67	DYS1	15.8	5.2	0.526
43,409	6.31	107	SER1	10.5	1.5	0.292
43,421	5.59	91	ERG6	2.2	14.1	0.408
44,174	7.32	56	YBR025C	13.1	6.0	0.684
44,682	4.99	72	TIF1	2.9	39.4	0.834
44,707	7.77	108	PGK1	23.7	165.7	0.897
44,707	7.77	109	PGK1	315.2	165.7	0.897
46,080	6.72	30	CAR2	15.4	NA[c]	0.495
46,383	8.52	53	IDP1	7.7	0.7	0.436
46,553	5.98	47	IDP2	32.4	NA[c]	0.197
46,679	6.39	50	ENO1	35.4	0.7	0.930
46,679	6.39	51	ENO1	6.6	0.7	0.930
46,679	6.39	52	ENO1	2.2	0.7	0.930
46,773	5.82	63	ENO2	15.5	289.1	0.960
46,773	5.82	64	ENO2	635.5	289.1	0.960
46,773	5.82	65	ENO2	93.0	289.1	0.960
46,773	5.82	66	ENO2	31.0	289.1	0.960
47,402	6.09	126	COR1	2.5	0.7	0.422
47,666	8.98	54	AAT2	11.7	6.0	0.338
48,364	5.25	73	WTM1	74.5	13.4	0.365
48,530	6.20	61	MET17	38.1	29.0	0.576
48,904	5.18	69	LYS9	16.2	3.7	0.463
48,987	4.90	153	SUP45	29.6	11.9	0.377
49,727	5.47	70	PRO2	13.6	5.2	0.297
49,912	9.27	62	TEF2	558.5	282.0	0.932
50,444	5.67	35	YDR190C	4.8	2.2	0.228
50,837	6.11	32	YEL047C	3.8	1.5	0.387
50,891	4.59	151	TUB2	11.2	7.4	0.404
51,547	6.80	27	LPD1	18.9	2.2	0.351
52,216	7.25	29	SHM2	19.7	7.4	0.722
52,859	5.54	37	YFR044C	30.2	6.7	0.442
53,798	5.19	71	HXK2	26.5	7.4	0.756
53,803	6.05	145	GYP6	4.4	0.7	0.147
54,403	5.29	39	ALD6	37.7	2.2	0.664
54,403	5.29	40	ALD6	6.6	2.2	0.664
54,502	6.20	31	ADE13	6.3	1.5	0.417
54,543	7.75	25	PYK1	225.3	101.8	0.965
54,543	7.75	26	PYK1	39.8	101.8	0.965
55,221	6.66	146	YEL071W	16.3	3.0	0.244
55,295	4.35	134	PDI1	66.2	14.1	0.589
55,364	5.98	24	GLK1	22.6	6.0	0.237
55,481	7.97	118	ATP1	21.6	2.2	0.637
55,886	6.47	28	CYS4	22.2	NA[c]	0.444
56,167	5.83	33	ARO8	14.3	3.0	0.324
56,167	5.83	34	ARO8	9.1	3.0	0.324
56,584	6.36	20	CYB2	18.9	NA[c]	0.259
57,366	5.53	60	FRS2	2.3	0.7	0.451
57,383	5.98	144	ZWF1	5.6	0.7	0.215
57,464	5.49	36	THR4	21.4	3.7	0.508
57,512	5.50	7	SRV2	6.5	NA[c]	0.260
57,727	4.92	152	VMA2	33.7	8.9	0.546
58,573	6.47	17	ACH1	4.4	1.5	0.327
58,573	6.47	18	ACH1	5.4	1.5	0.327
61,353	5.87	21	PDC1	6.5	200.7	0.962
61,353	5.87	22	PDC1	303.2	200.7	0.962
61,353	5.87	23	PDC1	16.3	200.7	0.962
61,649	5.54	38	CCT8	2.2	1.5	0.271

Continued

TABLE 1—*Continued*

Continued on following page

TABLE 1—*Continued*

Mol wt	pI	Spot no.	YPD gene name[a]	Protein abundance (10^3 copies/cell)	mRNA abundance (copies/cell)	Codon bias
61,902	6.21	101	PDC5	4.3	NA[c]	0.828
62,266	6.19	16	ICL1	20.1	NA[c]	0.327
62,862	8.02	19	ILV3	5.3	4.5	0.548
63,082	6.40	119	PGM2	2.2	3.0	0.402
64,335	5.77	5	PAB1	30.4	1.5	0.616
66,120	5.42	8	STI1	6.7	0.7	0.313
66,120	5.42	9	STI1	6.4	0.7	0.313
66,450	5.29	141	SSB2	7.0	NA[c]	0.880
66,450	5.29	142	SSB2	2.3	NA[c]	0.880
66,456	5.23	10	SSB1	64.5	79.5	0.907
66,456	5.23	11	SSB1	59.0	79.5	0.907
66,456	5.23	12	SSB1	13.7	79.5	0.907
68,397	5.82	82	LEU4	3.1	3.0	0.407
69,313	4.90	13	SSA2	24.3	18.6	0.892
69,313	4.90	14	SSA2	77.1	18.6	0.892
74,378	8.46	15	YKL029C	2.8	3.7	0.353
75,396	5.82	6	GRS1	5.5	7.4	0.500
85,720	6.25	1	MET6	2.0	NA[c]	0.772
85,720	6.25	2	MET6	10.9	NA[c]	0.772
85,720	6.25	3	MET6	1.4	NA[c]	0.772
93,276	6.11	131	EFT1	17.9	41.6	0.890
93,276	6.11	132	EFT1	5.7	41.6	0.890
102,064[e]	6.61[e]	94	ADE3	4.8	5.2	0.423
107,482[e]	5.33[e]	4	MCM3	2.7	NA[c]	0.240

[a] YPD gene names are available from the YPD website (39).
[b] NA, calculation could not be performed or was not available.
[c] mRNA data inconclusive or NA.
[d] No methionines in predicted ORF; therefore, protein concentration was not determined.
[e] Measured molecular weight or pI did not match theoretical molecular weight or pI.

Protein quantitation. [^{35}S]methionine-labeled gels were exposed to X-ray film overnight, and then the silver stain and film were used to excise 156 spots of varying intensities, molecular weights, and pIs. The excised spots were placed in 0.6-ml microcentrifuge tubes, and scintillation cocktail (100 μl) was added. The samples were vortexed and counted. In addition, two parallel gels were electroblotted to polyvinylidene difluoride membranes. The membranes were exposed to X-ray film, and four intense single spots were excised from each membrane and subjected to amino acid analysis. For these four spots, a mean of 209 ± 4 cpm/pmol of protein/methionine was found. This number was used to quantitate all remaining spots in conjunction with the number of methionines present in the protein.

To ensure that proteins were labeled to equilibrium, parallel 2D gels were prepared and run on yeast metabolically labeled for 1, 2, 6, or 18 h. The corresponding 156 spots were excised from each gel, and radioactivity was measured by liquid scintillation counting for each spot. Calculated protein levels were highly reproducible for all time points measured after 1 h.

Calculation of codon bias and predicted half-life. Codon bias values were extracted from the YPD spreadsheet (17). Protein half-lives were calculated based on the N-end rule (33). When the N-terminal processing was not known experimentally, it was predicted based on the affinity of methionine aminopeptidase (31).

RESULTS

Characteristics of proteome approach. Nearly every facet of proteome analysis hinges on the unambiguous identification of large numbers of expressed proteins in cells. Several techniques have been described previously for the identification of proteins separated by 2DE, including N-terminal and internal sequencing (1, 2), amino acid analysis (38), and more recently mass spectrometry (25). We utilized techniques based on mass spectrometry because they afford the highest levels of sensitivity and provide unambiguous identification. The specific procedure used is schematically illustrated in Fig. 1 and is based on three principles. First, proteins are removed from the gel by proteolytic in-gel digestion, and the resulting peptides are separated by on-line capillary high-performance liquid chromatography. Second, the eluting peptides are ionized and detected, and the specific peptide ions are selected and fragmented by the mass spectrometer. To achieve this, the mass spectrometer switches between the MS mode (for peptide mass identification) and the MS/MS mode (for peptide characterization and sequencing). Selected peptides are fragmented by a process called collision-induced dissociation (CID) to generate a tandem mass spectrum (MS/MS spectrum) that contains the peptide sequence information. Third, individual CID mass spectra are then compared by computer algorithms to predicted spectra from a sequence database. This results in the identification of the peptide and, by association, the protein(s) in the spot. Unambiguous protein identification is attained in a single analysis by the detection of multiple peptides derived from the same protein.

Protein identification. Yeast total cell protein lysate (40 μg), metabolically labeled with [^{35}S]methionine, was electrophoretically separated by isoelectric focusing in the first dimension and by SDS–10% polyacrylamide gel electrophoresis in the second dimension. Proteins were visualized by silver staining and by autoradiography. Of the more than 1,000 proteins visible by silver staining, 156 spots were excised from the gel and subjected to in-gel tryptic digestion, and the resulting peptides were analyzed and identified by microspray LC-MS/MS techniques as described above. The proteins in this study were all identified automatically by computer software with no human interpretation of mass spectra. They are indicated in Fig. 2 and detailed in Table 1.

The CID spectra shown in Fig. 3 indicate that the quality of the identification data generated was suitable for unambiguous protein identification. The spectra represent the amino acid sequences of tryptic peptides NSGDIVNLGSIAGR (Fig. 3A) and FAVGAFTDSLR (Fig. 3B). Both peptides were derived from protein S57593 (hypothetical protein YMR226C), which migrated to spot 114 (molecular weight, 29,156; pI, 6.59) in the 2D gel in Fig. 2. Five other peptides from the same analysis were also computer matched to the same protein sequence.

Protein and mRNA quantitation. For the 156 genes investigated, the protein expression levels ranged from 2,200 (PGM2) to 863,000 (TDH2/TDH3) copies/cell. The levels of mRNA for each of the genes identified were calculated from SAGE frequency tables (35). These tables contain the mRNA levels for 4,665 genes in yeast strain YPH499 grown to mid-log phase in YPD medium on glucose as a carbon source. In some instances, the mRNA levels could not be calculated for reasons stated in Materials and Methods. For the proteins analyzed in this study, mean transcript levels varied from 0.7 to 473 copies/cell.

Selection of the sample population for mRNA-protein expression level correlation. The protein spots selected for identification were selected from spots visible by silver staining in the 2D gel. An attempt was made not to include spots where overlap with other spots was readily apparent. The number of proteins identified was 156 (Table 1). Some proteins migrated to more than one spot (presumably due to differential protein processing or modifications), and protein levels from these spots were calculated by integrating the intensities of the different spots. The 156 protein spots analyzed represented the products of 128 different genes. Genes were excluded from the correlation analysis only if part of the data set was missing; i.e., genes were excluded if (i) no mRNA expression data were available for the protein or putative SAGE tags were ambiguous, (ii) the amino acid sequence did not contain methionine, (iii) more than a single protein was conclusively identified as

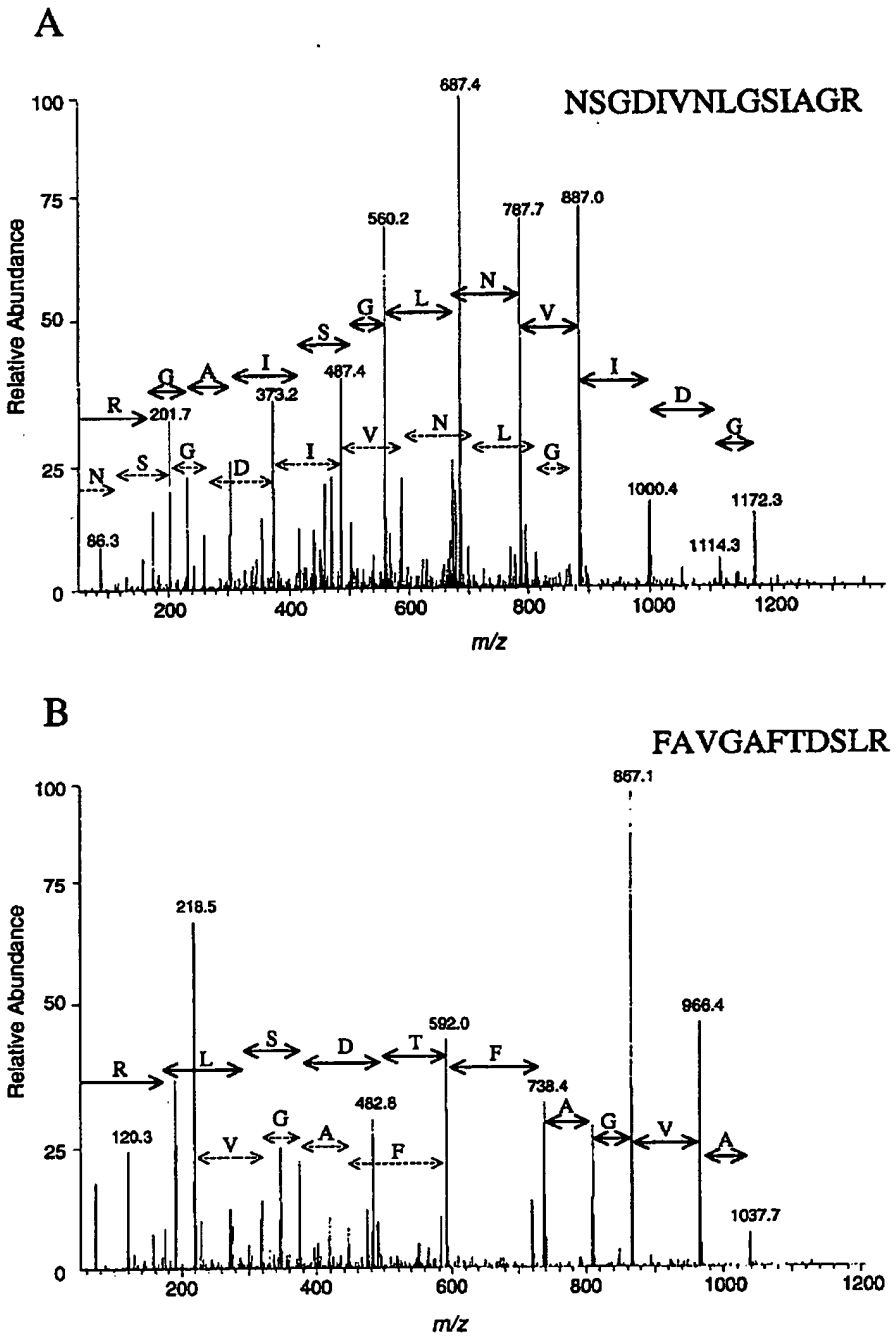


FIG. 3. Tandem mass (MS/MS) spectra resulting from analysis of a single spot on a 2D gel. The first quadrupole selected a single mass-to-charge ratio (*m/z*) of 687.2 (A) or 592.6 (B), while the collision cell was filled with argon gas, and a voltage which caused the peptide to undergo fragmentation by CID was applied. The third quadrupole scanned the mass range from 50 to 1,400 *m/z*. The computer program Sequest (8) was utilized to match MS/MS spectra to amino acid sequence by database searching. Both spectra matched peptides from the same protein, S57593 (yeast hypothetical protein YMR226C). Five other peptides from the same analysis were matched to the same protein.

migrating to the same gel spot, or (iv) the theoretical and observed pIs and molecular weights could not be reconciled. After these criteria were applied, the number of genes used in the correlation analysis was 106.

Codon bias and predicted half-lives. Codon bias is thought to be an indicator of protein expression, with highly expressed proteins having large codon bias values. The codon bias distribution for the entire set of more than 6,000 predicted yeast

gene ORFs is presented in Fig. 4A. The interval with the largest frequency of genes is between the codon bias values of 0.0 and 0.1. This segment contains more than 2,500 genes. The distribution of the codon bias values of the 128 different genes found in this study (all protein spots from Fig. 2) is shown in Fig. 4B, and protein half-lives (predicted from applying the N-end rule [33] to the experimentally determined or predicted protein N termini) are shown in Fig. 4C. No genes were identified with codon bias values less than 0.1 even though thousands of genes exist in this category. In addition, nearly all of the proteins identified had long predicted half-lives (greater than 30 h).

Correlation of mRNA and protein expression levels. The correlation between mRNA and protein levels of the genes selected as described above is shown in Fig. 5. For the entire group (106 genes) for which a complete data set was generated, there was a general trend of increased protein levels resulting from increased mRNA levels. The Pearson product moment correlation coefficient for the whole data set (106 genes) was 0.935. This number is highly biased by a small number of genes with very large protein and message levels. A more representative subset of the data is shown in the inset of Fig. 5. It shows genes for which the message level was below 10 copies/cell and includes 69% (73 of 106 genes) of the data used in the study. The Pearson product moment correlation coefficient for this data set was only 0.356. We also found that levels of protein expression coded for by mRNA with comparable abundance varied by as much as 30-fold and that the mRNA levels coding for proteins with comparable expression levels varied by as much as 20-fold.

The distortion of the correlation value induced by the uneven distribution of the data points along the x axis is further demonstrated by the analysis in Fig. 6. The 106 samples included in the study were ranked by protein abundance, and the Pearson product moment correlation coefficient was repeatedly calculated after including progressively more, and higher-abundance, proteins in each calculation. The correlation values remained relatively stable in the range of 0.1 to 0.4 if the lowest-expressed 40 to 95 proteins used in this study were included. However, the correlation value steadily climbed by the inclusion of each of the 11 very highly expressed proteins.

Correlation of protein and mRNA expression levels with codon bias. Codon bias is the propensity for a gene to utilize the same codon to encode an amino acid even though other codons would insert the identical amino acid in the growing polypeptide sequence. It is further thought that highly expressed proteins have large codon biases (3). To assess the value of codon bias for predicting mRNA and protein levels in exponentially growing yeast cells, we plotted the two experimental sets of data versus the codon bias (Fig. 7). The distribution patterns for both mRNA and protein levels with respect to codon bias were highly similar. There was high variability in the data within the codon bias range of 0.8 to 1.0. Although a large codon bias generally resulted in higher protein and message expression levels, codon bias did not appear to be predictive of either protein levels or mRNA levels in the cell.

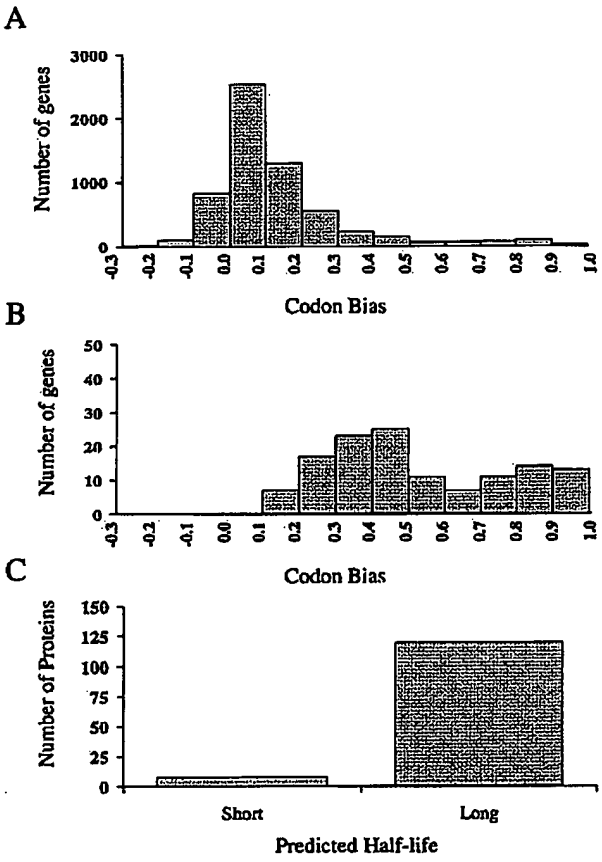


FIG. 4. Current proteome analysis technology utilizing 2DE without preenrichment samples mainly highly expressed and long-lived proteins. Genes encoding highly expressed proteins generally have large codon bias values. (A) Distribution of the yeast genome (more than 6,000 genes) based on codon bias. The interval with the largest frequency of genes is 0.0 to 0.1, with more than 2,500 genes. (B) Distribution of the genes from identified proteins in this study based on codon bias. No genes with codon bias values less than 0.1 were detected in this study. (C) Distribution of identified proteins in this study based on predicted half-life (estimated by N-end rule).

DISCUSSION

The desired end point for the description of a biological system is not the analysis of mRNA transcript levels alone but also the accurate measurement of protein expression levels and their respective activities. Quantitative analysis of global mRNA levels currently is a preferred method for the analysis of the state of cells and tissues (11). Several methods which either provide absolute mRNA abundance (34, 35) or relative mRNA levels in comparative analyses (20, 27) have been described elsewhere. The techniques are fast and exquisitely sensitive and can provide mRNA abundance for potentially any expressed gene. Measured mRNA levels are often implicitly or explicitly extrapolated to indicate the levels of activity of the corresponding protein in the cell. Quantitative analysis of protein expression levels (proteome analysis) is much more time-consuming because proteins are analyzed sequentially one by one and is not general because analyses are limited to the relatively highly expressed proteins. Proteome analysis does, however, provide types of data that are of critical importance for the description of the state of a biological system and that are not readily apparent from the sequence and the level of expression of the mRNA transcript. This study attempts to examine the relationship between mRNA and protein expression levels for a large number of expressed genes in cells representing the same state.

Limits in the sensitivity of current protein analysis technology precluded a completely random sampling of yeast proteins. We therefore based the study on those proteins visible by silver

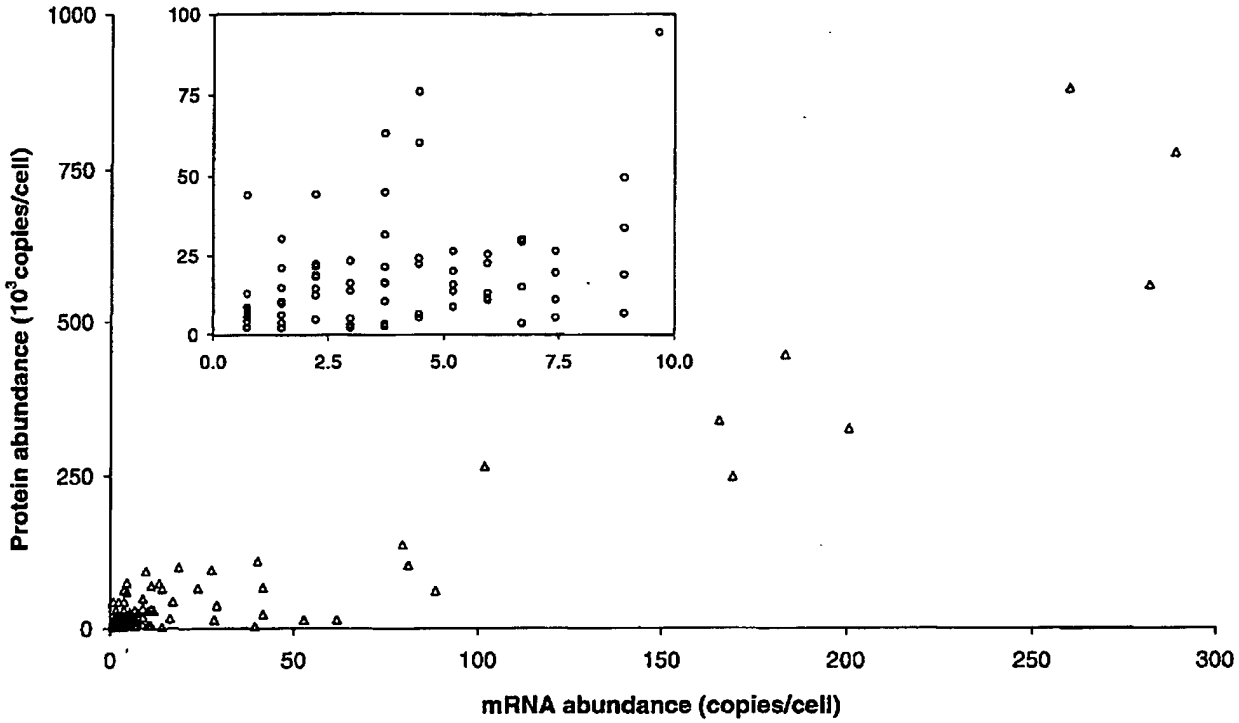


FIG. 5. Correlation between protein and mRNA levels for 106 genes in yeast growing at log phase with glucose as a carbon source. mRNA and protein levels were calculated as described in Materials and Methods. The data represent a population of genes with protein expression levels visible by silver staining on a 2D gel chosen to include the entire range of molecular weights, isoelectric focusing points, and staining intensities. The inset shows the low-end portion of the main figure. It contains 69% of the original data set. The Pearson product moment correlation for the entire data set was 0.935. The correlation for the inset containing 73 proteins (69%) was only 0.356.

staining on a 2D gel. Of the more than 1,000 visible spots, 156 were chosen to include the entire range of molecular weights, isoelectric focusing points, and staining intensities displayed on the 2D protein pattern. The genes identified in this study shared a number of properties. First, all of the proteins in this study had a codon bias of greater than 0.1 and 93% were greater than 0.2 (Fig. 4B). Second, with few exceptions, the proteins in this study had long predicted half-lives according to the N-end rule (Fig. 4C). Third, low-abundance proteins with regulatory functions such as transcription factors or protein kinases were not identified.

Because the population of proteins used in this study appears to be fairly homogeneous with respect to predicted half-life and codon bias, it might be expected that the correlation of the mRNA and protein expression levels would be stronger for this population than for a random sample of yeast proteins. We tested this assumption by evaluating the correlation value if different subsets of the available data were included in the calculation. The 106 proteins were ranked from lowest to highest protein expression level, and the trend in the correlation value was evaluated by progressively including more of the higher-abundance proteins in the calculation (Fig. 6). The correlation value when only the lower-abundance 40 to 93 proteins were examined was consistently between 0.1 and 0.4. If the 11 most abundant proteins were included, the correlation steadily increased to 0.94. We therefore expect that the correlation for all yeast proteins or for a random selection would be less than 0.4. The observed level of correlation between mRNA and protein expression levels suggests the importance of posttranslational mechanisms controlling gene expression. Such mechanisms include translational control (15) and control of protein half-life (33). Since these mechanisms are also active in higher eukaryotic cells, we speculate that there is no predictive correlation between steady-state levels of mRNA and those of protein in mammalian cells.

Like other large-scale analyses, the present study has several potential sources of error related to the methods used to determine mRNA and protein expression levels. The mRNA levels were calculated from frequency tables of SAGE data. This method is highly quantitative because it is based on actual sequencing of unique tags from each gene, and the number of times that a tag is represented is proportional to the number of mRNA molecules for a specific gene. This method has some limitations including the following: (i) the magnitude of the error in the measurement of mRNA levels is inversely proportional to the mRNA levels, (ii) SAGE tags from highly similar genes may not be distinguished and therefore are summed, (iii) some SAGE tags are from sequences in the 3′ untranslated region of the transcript, (iv) incomplete cleavage at the SAGE tag site by the restriction enzyme can result in two tags representing one mRNA, and (v) some transcripts actually do not generate a SAGE tag (34, 35).

For the SAGE method, the error associated with a value increases with a decreasing number of transcripts per cell. The conclusions drawn from this study are dependent on the quality of the mRNA levels from previously published data (35). Since more than 65% of the mRNA levels included in this study were calculated to 10 copies/cell or less (40% were less

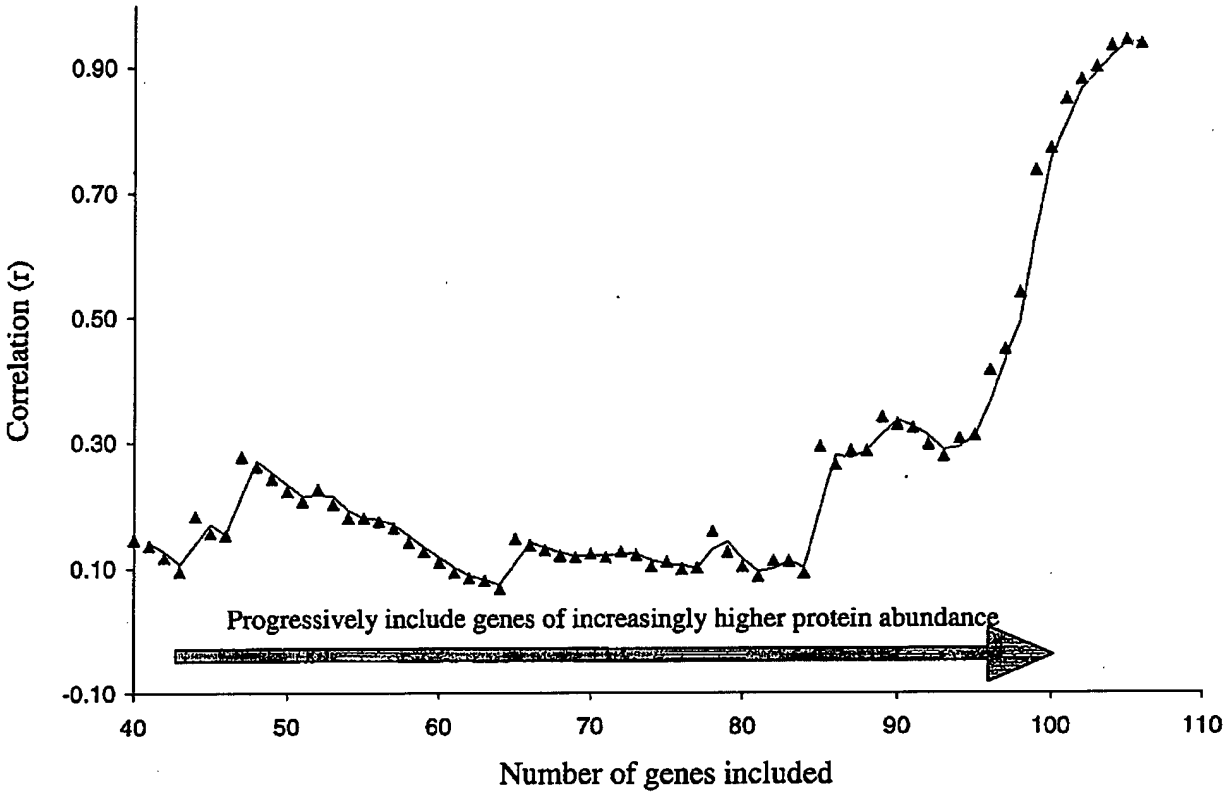


FIG. 6. Effect of highly abundant proteins on Pearson product moment correlation coefficient for mRNA and protein abundance in yeast. The set of 106 genes was ranked according to protein abundance, and the correlation value was calculated by including the 40 lowest-abundance genes and then progressively including the remaining 66 genes in order of abundance. The correlation value climbs as the final 11 highly abundant proteins are included.

than 4 copies/cell), the error associated with these values may be quite large. The mRNA levels were calculated from more than 20,000 transcripts. Assuming that the estimate of 15,000 mRNA molecules per cell is correct (16), this would mean that mRNA transcripts present at only a single copy per cell would be detected 72% of the time (35). The mRNA levels for each gene were carefully scrutinized, and only mRNA levels for which a high degree of confidence existed were included in the correlation value.

Protein abundance was determined by metabolic radiolabeling with [^{35}S]methionine. The calculation required knowledge of three variables: the number of methionines in the mature protein, the radioactivity contained in the protein, and the specific activity of the radiolabel normalized per methionine. The number of methionines per protein was determined from the amino acid sequence of the proteins identified by tandem mass spectrometry. For some proteins, it was not known whether the methionine of the nascent polypeptide was processed away. The N termini of those proteins were predicted based on the specificity of methionine aminopeptidase (31). If the N-terminal processing did not conform to the predicted specificity of processing enzymes, the calculation of the number of methionines would be affected. This discrepancy would affect most the quantitation of a protein with a very low number of methionines. The average number of calculated methionines per protein in this study was 7.2. We therefore expect the potential for erroneous protein quantitation due to unusual N-terminal processing to be small.

The amount of radioactivity contained in a single spot might be the sum of the radioactivity of comigrating proteins. Because protein identification was based on tandem mass spectrometric techniques, comigrating proteins could be identified. However, comigrating proteins were rarely detected in this study, most likely because relatively small amounts of total protein (40 μg) were initially loaded onto the gels, which resulted in highly focused spots containing generally 1 to 25 ng of protein. Because of the relatively small amount loaded, the concentrations of any potentially comigrating protein would likely be below the limit of detection of the mass spectrometry technique used in this study (1 to 5 ng) and below the limit of visualization by silver staining (1 to 5 ng). In the overwhelming majority of the samples analyzed, numerous peptides from a single protein were detected. It is assumed that any comigrating proteins were at levels too low to be detected and that their influence in the calculation would be small.

The specific activity of the radiolabel was determined by relating the precise amount of protein present in selected spots of a parallel gel, as determined by quantitative amino acid composition analysis, to the number of methionines present in the sequence of those proteins and the radioactivity determined by liquid scintillation counting. It is possible that the resulting number might be influenced by unavoidable losses inherent in the amino acid analysis procedure applied. Because four different proteins were utilized in the calculation and the experiment was done in duplicate, the specific activity calculated is thought to be highly accurate. Indeed, the specific

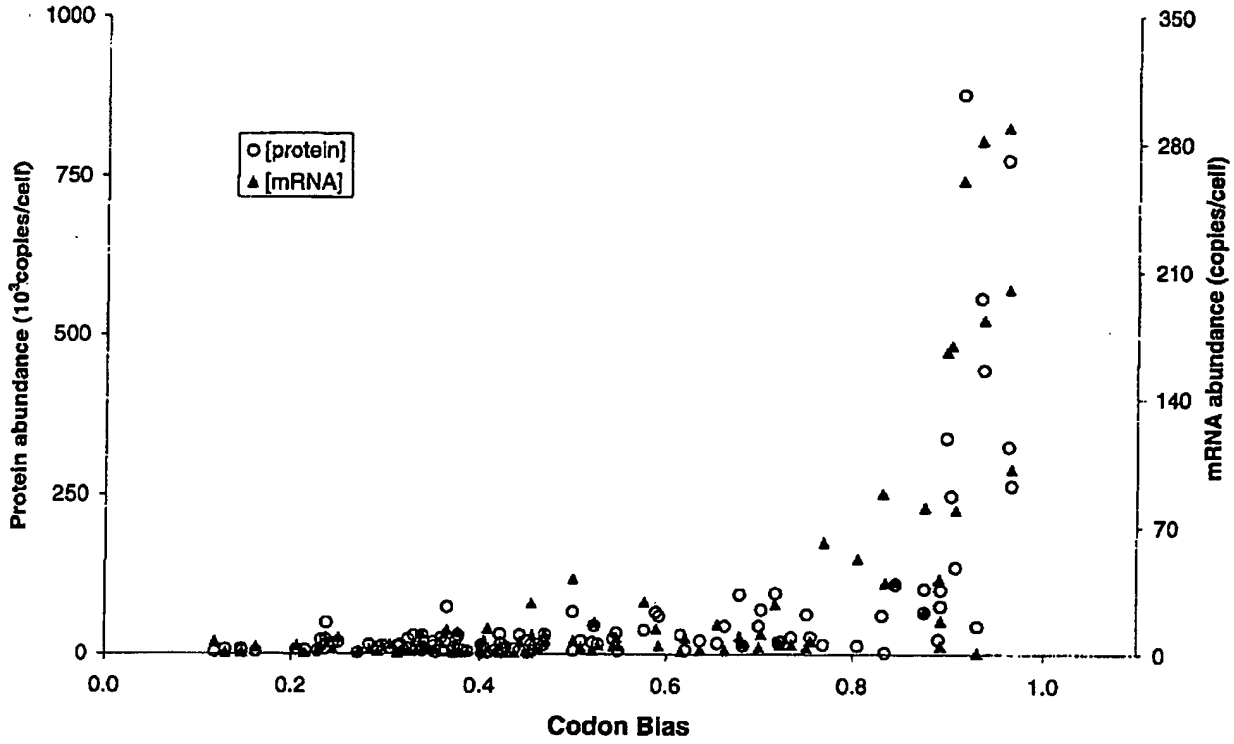


FIG. 7. Relationship between codon bias and protein and mRNA levels in this study. Yeast mRNA and protein expression levels were calculated as described in Materials and Methods. The data represent the same 106 genes as in Fig. 5.

activities calculated for each of the four proteins varied by less than 10%. Any inconsistencies in the calculation of the specific activity would result in differences in the absolute levels calculated but not in the relative numbers and would therefore not influence the correlation value determined.

The protein quantitative method used eliminates a number of potential errors inherent in previous methods for the quantitation of proteins separated by 2DE, such as preferential protein staining and bias caused by inequalities in the number of radiolabeled residues per protein. Any 2D gel-based method of quantitation is complicated by the fact that in some cases the translation products of the same mRNA migrated to different spots. One major reason is posttranslational modification or processing of the protein. Also, artifactual proteolysis during cell lysis and sample preparation can lead to multiple resolved forms of the protein. In such cases, the protein levels of spots coded for by the same mRNA were pooled. In addition, the existence of other spots coded for by the same mRNA that were not analyzed by mass spectrometry or that were below the limit of detection for silver staining cannot be ruled out. However, since this study is based on a class of highly expressed proteins, the presence of undetected minor spots below silver staining sensitivity corresponding to a protein analyzed in the study would generally cause a relatively small error in protein quantitation.

Codon bias is a measure of the propensity of an organism to selectively utilize certain codons which result in the incorporation of the same amino acid residue in a growing polypeptide chain. There are 61 possible codons that code for 20 amino acids. The larger the codon bias value, the smaller the number of codons that are used to encode the protein (19). It is thought that codon bias is a measure of protein abundance because highly expressed proteins generally have large codon bias values (3, 13).

Nearly all of the most highly expressed proteins had codon bias values of greater than 0.8. However, we detected a number of genes with high codon bias and relative low protein abundance (Fig. 7). For example, the expressed gene with both the second largest protein and mRNA levels in the study was ENO2_YEAST (775,000 and 289.1 copies/cell, respectively). ENO1_YEAST was also present in the gel at much lower protein and mRNA levels (44,200 and 0.7 copies/cell, respectively). The codon bias values for ENO2 and ENO1 are similar (0.96 and 0.93, respectively), but the expression of the two genes is differentially regulated. Specifically, ENO1_YEAST is glucose repressed (6) and was therefore present in low abundance under the conditions used. Other genes with large codon bias values that were not of high protein abundance in the gel include EFT1, TIF1, HXK2, GSP1, EGD2, SHM2, and TAL1. We conclude that merely determining the codon bias of a gene is not sufficient to predict its protein expression level.

Interestingly, codon bias appears to be an excellent indicator of the boundaries of current 2D gel proteome analysis technology. There are thousands of genes with expressed mRNA and likely expressed protein with codon bias values less than 0.1 (Fig. 4A). In this study, we detected none of them, and only a very small percentage of the genes detected in this study had codon bias values between 0.1 and 0.2 (Fig. 4B). Indeed, in every examined yeast proteome study (5, 7, 13, 28) where the combined total number of identified proteins is 300 to 400, this same observation is true. It is expected that for the more complex cells of higher eukaryotic organisms the detection of

low-abundance proteins would be even more challenging than for yeast. This indicates that highly abundant, long-lived proteins are overwhelmingly detected in proteome studies. If proteome analysis is to provide truly meaningful information about cellular processes, it must be able to penetrate to the level of regulatory proteins, including transcription factors and protein kinases. A promising approach is the use of narrow-range focusing gels with immobilized pH gradients (IPG) (23). This would allow for the loading of significantly more protein per pH unit covered and also provide increased resolution of proteins with similar electrophoretic mobilities. A standard pH gradient in an isoelectric focusing gel covers a 7-pH-unit range (pH 3 to 10) over 18 cm. A narrow-range focusing gel might expand the range to 0.5 pH units over 18 cm or more. This could potentially increase by more than 10-fold the number of proteins that can be detected. Clearly, current proteome technology is incapable of analyzing low-abundance regulatory proteins without employing an enrichment method for relatively low-abundance proteins. In conclusion, this study examined the relationship between yeast protein and message levels and revealed that transcript levels provide little predictive value with respect to the extent of protein expression.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation Science and Technology Center for Molecular Biotechnology, NIH grant T32HG00035-3, and a grant from Oxford Glycosciences.

We thank Jimmy Eng for expert computer programming, Garry Corthals and John R. Yates III for critical discussion, and Siavash Mohandesi for expert technical help.

REFERENCES

1. Aebersold, R. H., J. Leavitt, R. A. Saavedra, L. E. Hood, and S. B. Kent. 1987. Internal amino acid sequence analysis of proteins separated by one- or two-dimensional gel electrophoresis after in situ protease digestion on nitrocellulose. Proc. Natl. Acad. Sci. USA 84:6970–6974.
2. Aebersold, R. H., D. B. Teplow, L. E. Hood, and S. B. Kent. 1986. Electroblotting onto activated glass. High efficiency preparation of proteins from analytical sodium dodecyl sulfate-polyacrylamide gels for direct sequence analysis. Eur. J. Biochem. 261:4229–4238.
3. Bennetzen, J. L., and B. D. Hall. 1982. Codon selection in yeast. J. Biol. Chem. 257:3026–3031.
4. Boucherie, H., G. Dujardin, M. Kermorgant, C. Monribot, P. Slonimski, and M. Perrot. 1995. Two-dimensional protein map of Saccharomyces cerevisiae: construction of a gene-protein index. Yeast 11:601–613.
5. Boucherie, H., F. Sagliocco, R. Joubert, L Maillet, J. Labarre, and M. Perrot. 1996. Two-dimensional gel protein database of Saccharomyces cerevisiae. Electrophoresis 17:1683–1699.
6. Carmen, A. A., P. K. Brindle, C. S. Park, and M. J. Holland. 1995. Transcriptional regulation by an upstream repression sequence from the yeast enolase gene ENO1. Yeast 11:1031–1043.
7. Ducret, A., I. VanOostveen, J. K. Eng, J. R. Yates, and R. Aebersold. 1998. High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. Protein Sci. 7:706–719.
8. Eng, J., A. McCormack, and J. R. Yates. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom. 5:976–989.
9. Figeys, D., A. Ducret, J. R. Yates, and R. Aebersold. 1996. Protein identification by solid phase microextraction-capillary zone electrophoresis-microelectrospray-tandem mass spectrometry. Nat. Biotechnol. 14:1579–1583.
10. Figeys, D., I. VanOostveen, A. Ducret, and R. Aebersold. 1996. Protein identification by capillary zone electrophoresis/microelectrospray ionization-tandem mass spectrometry at the subfemtomole level. Anal. Chem. 68:1822–1828.
11. Fraser, C. M., and R. D. Fleischmann. 1997. Strategies for whole microbial genome sequencing and analysis. Electrophoresis 18:1207–1216.
12. Garrels, J. I., B. Futcher, R. Kobayashi, G. I. Latter, B. Schwender, T. Volpe, J. R. Warner, and C. S. McLaughlin. 1994. Protein identifications for a Saccharomyces cerevisiae protein database. Electrophoresis 15:1466–1486.
13. Garrels, J. I., C. S. McLaughlin, J. R. Warner, B. Futcher, G. I. Latter, R. Kobayashi, B. Schwender, T. Volpe, D. S. Anderson, F. Mesquita-Fuentes, and W. E. Payne. 1997. Proteome studies of Saccharomyces cerevisiae: identification and characterization of abundant proteins. Electrophoresis 18:1347–1360.
14. Gygi, S. P., and R. Aebersold. 1998. Absolute quantitation of 2-DE protein spots, p. 417–421. *In* A. J. Link (ed.), 2-D protocols for proteome analysis. Humana Press, Totowa, N.J.
15. Harford, J. B., and D. R. Morris. 1997. Post-transcriptional gene regulation. Wiley-Liss, Inc., New York, N.Y.
16. Hereford, L. M., and M. Rosbash. 1977. Number and distribution of polyadenylated RNA sequences in yeast. Cell 10:453–462.
17. Hodges, P. E., W. E. Payne, and J. I. Garrels. 1998. The Yeast Protein Database (YPD): a curated proteome database for Saccharomyces cerevisiae. Nucleic Acids Res. 26:68–72.
18. Klose, J., and U. Kobalz. 1995. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis 16:1034–1059.
19. Kurland, C. G. 1991. Codon bias and gene expression. FEBS Lett. 285:165–169.
20. Lashkari, D. A., J. L. DeRisi, J. H. McCusker, A. F. Namath, C. Gentile, S. Y. Hwang, P. O. Brown, and R. W. Davis. 1997. Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc. Natl. Acad. Sci. USA 94:13057–13062.
21. Liang, P., and A. B. Pardee. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967–971.
22. Link, A. J., L. G. Hays, E. B. Carmack, and J. R. Yates III. 1997. Identifying the major proteome components of Haemophilus influenzae type-strain NCTC 8143. Electrophoresis 18:1314–1334.
23. Nawrocki, A., M. R. Larsen, A. V. Podtelejnikov, O. N. Jensen, M. Mann, P. Roepstorff, A. Gorg, S. J. Fey, and P. M. Larsen. 1998. Correlation of acidic and basic carrier ampholyte and immobilized pH gradient two-dimensional gel electrophoresis patterns based on mass spectrometric protein identification. Electrophoresis 19:1024–1035.
24. O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. 250:4007–4021.
24a. OWL Protein Sequence Database. 2 August 1998, posting date. [Online.] http://bmbsgi11.leeds.ac.uk/bmb5dp/owl.html. [8 January 1999, last date accessed.]
25. Patterson, S. D., and R. Aebersold. 1995. Mass spectrometric approaches for the identification of gel-separated proteins. Electrophoresis 16:1791–1814.
26. Pennington, S. R., M. R. Wilkins, D. F. Hochstrasser, and M. J. Dunn. 1997. Proteome analysis: from protein characterization to biological function. Trends Cell Biol. 7:168–173.
27. Shalon, D., S. J. Smith, and P. O. Brown. 1996. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Res. 6:639–645.
28. Shevchenko, A., O. N. Jensen, A. V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, H. Boucherie, and M. Mann. 1996. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. USA 93:14440–14445.
29. Shevchenko, A., M. Wilm, O. Vorm, and M. Mann. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal. Chem. 68:850–858.
30. Sikorski, R. S., and P. Hieter. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics 122:19–27.
31. Tsunasawa, S., J. W. Stewart, and F. Sherman. 1985. Amino-terminal processing of mutant forms of yeast iso-1-cytochrome c. The specificities of methionine aminopeptidase and acetyltransferase. J. Biol. Chem. 260:5382–5391.
32. Urlinger, S., K. Kuchler, T. H. Meyer, S. Uebel, and R. Tamp'e. 1997. Intracellular location, complex formation, and function of the transporter associated with antigen processing in yeast. Eur. J. Biochem. 245:266–272.
33. Varshavsky, A. 1996. The N-end rule: functions, mysteries, uses. Proc. Natl. Acad. Sci. USA 93:12142–12149.
34. Velculescu, V. E., L. Zhang, B. Vogelstein, and K. W. Kinzler. 1995. Serial analysis of gene expression. Science 270:484–487.
35. Velculescu, V. E., L. Zhang, W. Zhou, J. Vogelstein, M. A. Basrai, D. E. Bassett, Jr., P. Hieter, B. Vogelstein, and K. W. Kinzler. 1997. Characterization of the yeast transcriptome. Cell 88:243–251.
36. Wilkins, M. R., K. L. Williams, R. D. Appel, and D. F. Hochstrasser. 1997. Proteome research: new frontiers in functional genomics. Springer-Verlag, Berlin, Germany.
37. Wilm, M., A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, and M. Mann. 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379:466–469.
38. Yan, J. X., M. R. Wilkins, K. Ou, A. A. Gooley, K. L. Williams, J. C. Sanchez, O. Golaz, C. Pasquali, and D. F. Hochstrasser. 1996. Large-scale amino-acid analysis for proteome studies. J. Chromatogr. A 736:291–302.
39. YPD Website. 6 March 1998, revision date. [Online.] Proteome, Inc. http://www.proteome.com/YPDhome.html. [8 January 1999, last date accessed.]

EXHIBIT 10

(12) **United States Patent**
Sundick et al.

(10) **Patent No.: US 6,737,522 B2**
(45) **Date of Patent: *May 18, 2004**

(54) **CHICKEN INTERLEUKIN-15 AND USES THEREOF**

(75) Inventors: **Roy S. Sundick**, Farmington Hills, MI (US); **Lily A. Jones**, Grosse Point Park, MI (US); **David I. Smith**, Rochester, MN (US)

(73) Assignee: **Wayne State University**, Detroit, MI (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **09/907,539**

(22) Filed: **Jul. 17, 2001**

(65) **Prior Publication Data**

US 2002/0150554 A1 Oct. 17, 2002

Related U.S. Application Data

(62) Division of application No. 09/368,613, filed on Aug. 4, 1999, now Pat. No. 6,287,554, and a division of application No. 08/729,004, filed on Oct. 10, 1996, now Pat. No. 6,190,901.

(60) Provisional application No. 60/005,682, filed on Oct. 17, 1995.

(51) **Int. Cl.**[7] **C07H 21/04**; C12N 15/00; C12N 15/63

(52) **U.S. Cl.** **536/23.51**; 536/23.1; 435/320.1

(58) **Field of Search** 536/23.11, 23.51; 435/320.11, 455.1, 471

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,490,289	A	12/1984	Stern
4,992,367	A	2/1991	Cullen
5,106,617	A	4/1992	Federicksen et al.
5,149,788	A	9/1992	Bailon et al.
5,162,503	A	11/1992	Bailon et al.
5,455,337	A	10/1995	Devos et al.
5,536,657	A	7/1996	Chua et al.
5,597,901	A	1/1997	Stern

FOREIGN PATENT DOCUMENTS

EP	0 293 180	11/1988	..... C12P/21/02
EP	0 473 696 B1	7/1995	
WO	WO95/27722	10/1995	
WO	WO96/04306	2/1996	
WO	WO96/37223	2/1996	
ZA	9 402 636	11/1994	

OTHER PUBLICATIONS

IL–15 Stimulates The Expansion of Aids Virus–specific CTL[1] (The Journal of Immunology, 1996, 157: 3681–3687.

Studies Evaluating The Antitumor Activity And Toxicity of Interleukin–15, a New T Cell Growth Factor: Comparison With Interleukin–2 (Immunex Research And Development Corporation, Seattle, Washington 98101 And Oncology And Immunology Research Section, Lederle Laboratories, Pearl River, New York 10965, 1965.

Dialog(R)File 149:iac(sm) Health & Wellness Db(sm) (C) 1997 Info Access Co. (Il–15 Boosts Systemic Immunity Against B7–1 Gene Transfected Breast Cancer Cells (Interleukin 15).

Lillehoj et al., "Avian Interferon and Interleukin–2. A Review by Comparison With Mammalian Homologues" *Poultry Science Research*, vol. 4, No. 2, pp. 67–85, 1992.

Interleukin–15, A New T Cell Growth Factor: Comparison with Interleukin–2 (Immunex Research and Development Corporation, Seattle, Washington 98101 and Oncology and Immunology Research Section, Lederle Laboratories, Pearl River, New York 10965) 1995.

Sin et al., J. Virol. vol. 73, 501–509, 1999.

Rodolfo and Colombo, Methods, vol. 19, 114–120, 1999.

Openshaw and Hussell, Dev. Biol. Stand., vol. 92, 179–185, 1998.

Armitage et al., "IL–15 Has Stimulatory Activity for the Induction of B Cell Proliferation and Differentiation" Journal of Immunology, vol. 154, No. 2, pp. 483–490, 1995.

Gini et al., "IL–15 A Novel T Cell Growth Factor That Shares Activities and Receptor Components With IL–2" Journal of Leukocyte Biology, vol. 57, No. 5, pp. 763–766, 1995.

Staeheli et al., "Cytokines of Birds: Conserved Functions–A Largely Different Look" Journal of Interferon and Cytokine, vol. 21, No. 12, pp. 993–1010, 2001.

Primary Examiner—David Guzo
Assistant Examiner—Quang Nguyen
(74) *Attorney, Agent, or Firm*—Darby & Darby

(57) **ABSTRACT**

The present invention pertains to isolated DNA encoding avian interleukin-15 and to purified interleukin-15 polypeptides.

27 Claims, 3 Drawing Sheets

FIG. 1A

T7 end of the pCDNA1 vector:

5'-TGCTTGGTACCGAGCTCGGATCCACTAGTAACGCCCGCCAGTGTGCTCTAAAG-

Noncoding Segment of cDNA: *CAGATAACTGGGACACTGCC

Coding Region of First Open Region Frame (IL-15):

ATGATGTGCAAAGTACTGATCTTTGGCTGTATTTCGGTAGCAACGCTAATG

ACTACAGCTTATGGAGCATCTCTATCATCAGCAAAAAGGAAACCTCTTCAA

ACATTAATAAAGGATTTAGAAATATTGGAAAATATCAAGAACAAGATTCAT

CTCGAGCTCTACACACCAACTGAGACCCAGGAGTGCACCCAGCAAACTCTG

CAGTGTTACCTGGGAGAAGTGGTTACTCTGAAGAAAGAAACTGAAGATGAC

ACTGAAATTAAAGAAGAATTTGTAACTGCTATTCAAAATATCGAAAAGAAC

CTCAAGAGTCTTACGGGTCTAAATCACACCGGAAGTGAATGCAAGATCTGT

GAAGCTAACAACAAGAAAAAATTTCCTGATTTTCTCCATGAACTGACCAAC

TTTGTGAGATATCTGCAAAAA

A ——————————————————————————————— A

FIG. 1B

A ——— A

Sequence of Remaining Insert cDNA:

TAAGCAACTAATCATTTTTATTTTACTGCTATGTTATTTATTTAATTATTT

AATTACAGATAATTTATATATTTTATCCCGTGGCTAACTAATCTGCTGTCC

ATTCTGGGACCACTGTATGCTCTTAGTCTGGGTGATATGACGTCTGTTCTA

AGATCATATTTGATCCTTTCTGTAACCTACGGGCTCAAAATGTACGTTGGA

AAACTGATTGATTCTCACTTTGTCGGTAAAGTGATATGTGTTTACTGAAAG

AATTTTTAAAAGTCACTTCTAGATGACATTTAATAAATTTCAG#

Sp6 end of the pcDNA1 vector:

CTTTAGAGCACACTGGCGGCCCNTCGAGCATGCATCTAGAGGGCC-3'

* beginning of cDNA

end of cDNA

FIG. 2

chicken IL-15 precursor, 143 amino acids

MMCKVLIFGCISVATLMTTAYGASLSSAKRKPLQTLIKDLEILENIKNKI
HLELYTPTETQECTQQTLQCYLGEVVTLKKETEDDTEIKEEFVTAIQNIE
KNLKSLTGLNHTGSECKICEANNKKFPDFLHELTNFVRYLQK

CHICKEN INTERLEUKIN-15 AND USES THEREOF

This is a divisional of U.S. patent application Ser. No. 09/368,613 filed Aug. 4, 1999, which issued as U.S. Pat. No. 6,287,554 and is a divisional of U.S. patent application Ser. No. 08/729,004, filed on Oct. 10, 1996 and now issued as U.S. Pat. No. 6,190,901. U.S. patent application Ser. No. 08/729,004 claims priority to U.S. provisional patent application Serial No. 60/005,682 filed on Oct. 17, 1995. Each of these prior applications is hereby incorporated herein by reference, in its entirety.

FIELD OF INVENTION

The present invention pertains to isolated genes encoding avian interleukin-15 and to purified interleukin-15 polypeptides.

BACKGROUND OF THE INVENTION

Most chickens produced in developed countries for consumption and egg-laying (at least 10 billion per year) are vaccinated to protect them against Marek's disease. All of the egg-laying chickens and breeder stocks are also vaccinated with Newcastle Disease Virus, Infectious Bursal Disease Virus, Infectious Bronchitis Virus, Fowlpox Virus and Coccidial vaccines. For optimal protection, Marek's vaccination is performed either at or before hatching. One obstacle to the development of efficacious pre-hatching and at-hatching vaccination regimens is that the embryonic and newly hatched avian immune system is not fully developed and cannot mount as effective an immune response to the immunogen as at 2–3 weeks after hatching. Thus, there is a need in the art for agents and compositions that enhance the effectiveness of pre- and post-hatching avian vaccines.

Interleukin-2 and interleukin-15 are related cytokines that stimulate the activity and proliferation of T cells in mammals. Though IL-2 and IL-15 both interact with the β and γ chains of the IL-2 receptor, and may share some elements of tertiary structure, the two polypeptides are not homologous and represent distinct gene products.

The genes encoding IL-15 from several different mammalian species share a high degree of homology. For example, human and simian IL-15 share 97% amino acid homology. By contrast, chicken IL-15, which is the subject of the present invention, shares only 25% amino acid identity with mammalian IL-15. Another distinguishing characteristic of chicken IL-15 is that it (and not the mammalian forms) is produced by mitogen-activated spleen cells. Accordingly, the discovery of chicken IL-15 and the finding that it possesses T cell-stimulatory activity provide a novel reagent for vaccine augmentation in avian species. Without wishing to be bound by theory, the bioactivity of mammalian IL-15 in stimulating skeletal muscle development suggests that avian IL-15s are also useful in stimulating growth in avian species.

SUMMARY OF THE INVENTION

The present invention provides isolated and purified DNA encoding avian interleukin-15 (IL-15), as well as cloning and expression vectors comprising IL-15 DNA and cells transformed with IL-15-encoding vectors. Avian species from which IL-15 may be derived include without limitation chicken, turkey, duck, goose, quail and pheasant.

The invention also provides isolated and purified avian IL-15 polypeptide, the native secreted or mature form of which has a molecular mass of about 14 kDa, an isoelectric point of about 6.57, a net charge of −2, and a hydrophilicity index of 0.278, and which has the ability to stimulate mitogen-activated avian T cells and to promote the growth of other cell types. IL-15 according to the present invention may be obtained from native or recombinant sources.

Also encompassed by the invention are sequence-conservative and function-conservative variants of avian IL-15 DNA and IL-15 polypeptides, including, for example, a bioactive IL-15 sequence or sub-fragment that is fused in-frame to a purification sequence.

In another aspect, the invention provides a method for enhancing an immune response in fowl to an immunogen, which is achieved by administering the immunogen before, after, or substantially simultaneously with avian IL-15 in an amount effective to enhance the immune response.

In yet another aspect, the invention provides a vaccine for inducing an immune response in fowl to an immunogen, comprising the immunogen and an effective amount of avian interleukin-15 for immune response enhancement. The immunogen may be derived, for example, from avian pathogens such as Marek's Disease Virus, Newcastle Disease Virus, Infectious Bursal Disease Virus, Infectious Bronchitis Virus, and the like.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is an illustration of an 845 nt sequence including 747 nt of cDNA sequence encoding chicken interleukin-15 (IL-15) SEQ ID NO:1.

FIG. **2** is an illustration of a 143-amino acid sequence corresponding to the chicken interleukin-15 precursor polypeptide (SEQ ID NO:2).

DETAILED DESCRIPTION OF THE INVENTION

All patent applications, patents, and literature references cited in this specification are hereby incorporated by reference in their entirety. In case of conflict, the present description, including definitions, will control.

The present invention encompasses interleukin-15 (IL-15) from avian species. The invention provides isolated and purified nucleic acids encoding avian IL-15, as well as IL-15 polypeptides purified from either native or recombinant sources. Avian IL-15 produced according to the present invention may be used in commercial fowl cultivation to promote growth and to enhance the efficacy of avian vaccines.

Nucleic Acids, Vectors, Transformants

The sequence of the cDNA encoding chicken IL-15 is shown in FIG. **1** (SEQ ID NO:1), and the predicted amino acid sequence of chicken IL-15 is shown in FIG. **2** (SEQ ID NO:2). The designation of this avian polypeptide as IL-15 is based on partial amino acid sequence homology to mammalian IL-15 and the ability of the polypeptide to stimulate mitogen-activated T cells (see below). Furthermore, without wishing to be bound by theory, it is predicted that avian IL-15 polypeptides also exhibit one or more of the following bioactivities: activation of NK (natural killer) cells, stimulation of B-Cell maturation, proliferation of mast cells, and interaction with the beta and gamma subunits of the IL-2 receptor.

Because of the degeneracy of the genetic code (i.e., multiple codons encode certain amino acids), DNA sequences other than that shown in FIG. **1** can also encode the chicken IL-15 amino acid sequences shown in FIG. **2**.

Such other DNAs include those containing "sequence-conservative" variations in which a change in one or more nucleotides in a given codon results in no alteration in the amino acid encoded at that position. Furthermore, a given amino acid residue in a polypeptide can often be changed without altering the overall conformation and function of the native polypeptide. Such "function-conservative" variants include, but are not limited to, replacement of an amino acid with one having similar physico-chemical properties, such as, for example, acidic, basic, hydrophobic, and the like (e.g., replacement of lysine with arginine, aspartate with glutamate, or glycine with alanine). In addition, amino acid sequences may be added or deleted without destroying the bioactivity of the molecule. For example, additional amino acid sequences may be added at either amino- or carboxy-terminal ends to serve as purification tags, (i.e., to allow one-step purification of the protein, after which they may be chemically or enzymatically removed). Alternatively, the additional sequences may confer an additional cell-surface binding site or otherwise alter the target cell specificity of IL-15.

The chicken IL-15 cDNAs within the scope of the present invention are those of FIG. **1**, sequence-conservative variant DNAs, DNA sequences encoding function-conservative variant polypeptides, and combinations thereof. The invention encompasses fragments of avian interleukin-15 that exhibit a useful degree of bioactivity, either alone or in combination with other sequences or components. As explained below, it is well within the ordinary skill in the art to predictively manipulate the sequence of IL-15 and establish whether a given avian IL-15 variant possesses an appropriate stability and bioactivity for a given application. This can be achieved by expressing and purifying the variant IL-15 polypeptide in a recombinant system and assaying its T-cell stimulatory activity and/or growth-promoting activity in cell culture and in animals, followed by testing in the application.

The present invention also encompasses IL-15 DNAs (and polypeptides) derived from other avian species, including without limitation ducks, turkeys, pheasants, quail and geese. Avian IL-15 homologues of the chicken sequence shown in FIG. **1** are easily identified by screening cDNA or genomic libraries to identify clones that hybridize to probes comprising all or part of the sequence of FIG. **1**. Alternatively, expression libraries may be screened using antibodies that recognize chicken IL-15. Without wishing to be bound by theory, it is anticipated that IL-15 genes from other avian species will share at least about 70% homology with the chicken IL-15 gene. Also within the scope of the invention are DNAs that encode chicken homologues of IL-15, defined as DNA encoding polypeptides that share at least about 25% amino acid identity with chicken IL-15.

Generally, nucleic acid manipulations according to the present invention use methods that are well known in the art, such as those as disclosed in, for example, *Molecular Cloning, A Laboratory Manual* (2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor), or *Current Protocols in Molecular Biology* (Eds. Aufubel, Brent, Kingston, More, Feidman, Smith and Stuhl, Greene Publ. Assoc., Wiley-Interscience, NY, N.Y., 1992).

The present invention encompasses cDNA and RNA sequences and sense and antisense sequences. The invention also encompasses genomic avian IL-15 polypeptide DNA sequences and flanking sequences, including, but not limited to, regulatory sequences. Nucleic acid sequences encoding avian IL-15 polypeptide(s) may also be associated with heterologous sequences, including promoters, enhancers, response elements, signal sequences, polyadenylation sequences, introns, 5'- and 3'- noncoding regions, and the like. Transcriptional regulatory elements that may be operably linked to avian IL-15 polypeptide DNA sequence(s) include without limitation those that have the ability to direct the expression of genes derived from prokaryotic cells, eukaryotic cells, viruses of prokaryotic cells, viruses of eukaryotic cells, and any combination thereof. Other useful heterologous sequences are known to those skilled in the art.

The nucleic acids of the present invention can be modified by methods known to those skilled in the art to alter their stability, solubility, binding affinity, and specificity. For example, the sequences can be selectively methylated. The nucleic acid sequences of the present invention may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

The present invention also provides vectors that include nucleic acids encoding the avian IL-15 polypeptide(s). Such vectors include, for example, plasmid vectors for expression in a variety of eukaryotic and prokaryotic hosts. Preferably, vectors also include a promoter operably linked to the avian IL-15 polypeptide encoding portion. The encoded avian IL-15 polypeptide(s) may be expressed by using any suitable vectors and host cells as explained herein or otherwise known to those skilled in the art.

Vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host such as, for example, antibiotic resistance, and one or more expression cassettes. The inserted coding sequences may be synthesized, isolated from natural sources, prepared as hybrids, or the like. Ligation of the coding sequences to the transcriptional regulatory sequences may be achieved by methods known to those skilled in the art. Suitable host cells may be transformed/transfected/infected by any suitable method including electroporation, $CaCl_2$- or liposome-mediated DNA uptake, fungal infection, microinjection, microprojectile, or the like.

Suitable vectors for use in practicing the present invention include without limitation YEp352, pcDNAI (In Vitrogen, San Diego, Calif.), pRc/CMV (InVitrogen), and pSFV1 (GIBCO/BRL, Gaithersburg, Md.). One preferred vector for use in the invention is pSFV1. Suitable host cells include *E. Coli*, yeast, COS cells, PC12 cells, CHO cells, GH4C1 cells, BHK-21 cells, and amphibian melanophore cells. BHK-21 cells are a preferred host cell line for use in practicing the present invention.

Nucleic acids encoding avian IL-15 polypeptide(s) may also be introduced into cells by recombination events. For example, such a sequence can be microinjected into a cell, effecting homologous recombination at the site of an endogenous gene encoding the polypeptide, an analog or pseudogene thereof, or a sequence with substantial identity to an avian IL-15 polypeptide-encoding gene. Other recombination-based methods such as non-homologous recombinations, and deletion of endogenous gene by homologous recombination, especially in pluripotent cells, may also be used.

IL-15 Polypeptides

The chicken IL-15 gene (the cDNA of which is shown in FIG. **1**) encodes a polypeptide of 143 amino acids (FIG. **2**). Without wishing to be bound by theory, by comparison with simian IL-15, and by use of an accepted procedure to predict signal peptidase cleavage sites (Von Heijne, *Nuc.Acids Res.*, 14:4683, 1986), it is predicted that an aminoterminal leader sequence of about 22 amino acids (secretion signal peptide)

is cleaved from the primary translation product to produce mature IL-15. The predicted mature sequence of 121 amino acids is further characterized by a predicted molecular weight of 13,971 daltons; an isoelectric point of 6.57; four cysteine residues (at amino acids numbers 63, 70, 116, and 119 in the precursor IL-15 shown in FIG. 2) that correspond to four cysteines conserved among human, mouse, and monkey IL-15 and that are believed to participate in intramolecular disulfide bonding; and one consensus site for N-linked glycosylation (at asparagine 110 of the sequence shown in FIG. 2) which corresponds to a similar site in human IL-15.

Purification of IL-15 from natural or recombinant sources may be achieved by methods well-known in the art, including without limitation ion-exchange chromatography, reverse-phase chromatography on C4 columns, gel filtration, isoelectric focusing, affinity chromatography, immunoaffinity chromatography, and the like. In a preferred embodiment, large quantities of bioactive IL-15 may be obtained by constructing a recombinant DNA sequence comprising the coding region for IL-15 fused in frame to a sequence encoding 6 C-terminal histidine residues in the pSFV1 replicon (GIBCO/BRL). mRNA encoded by this plasmid is synthesized using techniques well-known to those skilled in the art and introduced into BHK-21 cells by electroporation. The cells synthesize and secrete mature glycosylated IL-15 polypeptides containing 6 C-terminal histidines. The modified IL-15 polypeptides are easily purified from the cell supernatant by affinity chromatography using a histidine-binding resin (His-bind, Novagen, Madison, Wis.).

Avian IL-15 polypeptides isolated from any source can be modified by methods known in the art. For example, avian IL-15 may be phosphorylated or dephosphorylated, glycosylated or deglycosylated, and the like. Especially useful are modifications that alter avian IL-15 solubility, stability, and binding specificity and affinity.

Anti-IL-15 Antibodies

The present invention encompasses antibodies that are specific for avian IL-15 polypeptides identified as described above. The antibodies may be polyclonal or monoclonal, and may discriminate avian IL-15s from different species, identify functional domains, and the like. Such antibodies are conveniently made using the methods and compositions disclosed in Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988, other references cited herein, as well as immunological and hybridoma technologies known to those in the art. Where natural or synthetic avian IL-15-derived peptides are used to induce an avian IL-15-specific immune response, the peptides may be conveniently coupled to a suitable carrier such as KLH and administered in a suitable adjuvant such as Freund's. Preferably, selected peptides are coupled to a lysine core carrier substantially according to the methods of Tam (1988) *Proc. Natl. Acad. Sci. USA*, 85:5409–5413. The resulting antibodies may be modified to a monovalent form e.g. Fab, FAB', or FV. Anti-idiotypic antibodies, especially internal imaging anti-idiotypic antibodies, may also be prepared using known methods.

In one embodiment, purified avian IL-15 is used to immunize mice, after which their spleens are removed, and splenocytes used to form cell hybrids with myeloma cells to obtain clones of antibody-secreting cells according to techniques that are standard in the art. The resulting monoclonal antibodies secreted by such cells are screened using in vitro assays for the following activities: binding to avian IL-15, inhibiting the receptor-binding activity of IL-15, and inhibiting the T-cell stimulatory activity of IL-15.

Anti-avian IL-15 antibodies may be used to identify and quantify avian IL-15, using immunoassays such as ELISA, RIA, and the like. Anti-avian IL-15 antibodies may also be used to immunodeplete extracts of avian IL-15. In addition, these antibodies can be used to identify, isolate, and purify avian IL-15s from different sources, and to perform subcellular and histochemical localization studies.

Applications

Avian IL-15 produced according to the present invention can be used beneficially in homologous or heterologous avian species, for example, to stimulate activated T-cells (Grabstein et al., *Science,* 264:965, 1994) and B-cells (Armitage et al., *J. Immunol.,* 154:483, 1995) and/or to promote the growth of non-immune cells, such as, for example, muscle cells (Quinn et al., *Endocrinol.* 136:3669, 1995).

Vaccines

The present invention encompasses methods and compositions for enhancing the efficacy of an immune response in avian species. In this embodiment, avian IL-15 is used in conjunction with an immunogen for which it is desired to elicit an immune response. For example, in avian vaccines, such as those against Marek's disease, Newcastle Disease Virus, and other pathogens such as Infectious Bursal Disease Virus and Infectious Bronchitis Virus, it is desirable to include avian IL-15 in the vaccine to enhance the magnitude and quality of the immune response. For this purpose, IL-15 purified from native or recombinant sources as described above is included in the vaccine formulation at a concentration ranging from about 0.01 μg to about 1.0 μg per vaccine per chicken.

IL-15 may be administered in conjunction with a live (i.e., replicating) vaccine or a non-replicating vaccine. Non-limiting examples of replicating vaccines are those comprising native or recombinant viruses or bacteria, such as modified turkey herpesvirus or modified fowlpox virus. Non-limiting examples of non-replicating vaccines are those comprising killed or inactivated viruses or other microorganisms, or crude or purified antigens derived from native, recombinant, or synthetic sources, such as, for example, coccidial vaccines. Commercial sources for avian vaccines include without limitation: Rhone Merieux Laboratoire-IFFA (Lyon, France); Intervet International BV (Boxmeer, The Netherlands); Mallinckrodt Veterinary; Solvay Animal Health (Mendota Heights, Minn.); Hoechst-Roussel (Knoxville, Tenn.); and Nippon Zeon Co., Ltd. (Kawasaki-Kiu, Japan).

In one embodiment, the gene encoding IL-15 is incorporated into a recombinant virus, which is then formulated into a live vaccine. The IL-15 gene is incorporated into the virus so that its expression is controlled by an appropriate promoter. Administration of the vaccine results in the expression of bioactive IL-15 in close temporal and spatial proximity to the desired immune response, thus enhancing the vaccine's efficacy.

IL-15 may be administered to birds as part of a vaccine formulation either before or after hatching, preferably before hatching, using methods known in the art such as those described in U.S. Pat. Nos. 5,034,513 and 5,028,421.

Growth Promotion

The present invention provides methods and compositions for enhancing the growth of avian species for medical and/or commercial purposes. In this embodiment, IL-15 is administered to birds using any appropriate mode of administration. For growth promotion, IL-15 is administered in amounts ranging from about 0.25 μg/kg/day to about 25 μg/kg/day. It will be understood that the required amount of

IL-15 can be determined by routine experimentation well-known in the art, such as by establishing a matrix of dosages and frequencies and comparing a group of experimental units or subjects to each point in the matrix.

According to the present invention, native or recombinant avian IL-15 may be formulated with a physiologically acceptable carrier, such as, for example, phosphate buffered saline or deionized water. The formulation may also contain excipients, including lubricant(s), plasticizer(s), colorant(s), absorption enhancer(s), bactericide(s), and the like that are well-known in the art. The IL-15 polypeptide of the invention may be administered by any effective means, including without limitation intravenous, subcutaneous, intramuscular, transmucosal, topical, or oral routes. For subcutaneous administration, for example, the dosage form may consist of IL-15 in sterile physiological saline. For oral administration, IL-15, with or without excipients, may be micro- or macro-encapsulated in, e.g., liposomes and microspheres. Dermal patches (or other slow-release dosage forms) may also be used.

The following examples are intended to further illustrate the invention without limiting its scope thereof.

EXAMPLE 1

Cloning of the Chicken IL-15 Gene

To clone chicken IL-15, a chicken spleen cell cDNA library derived from spleen cells that had been activated with concanavalin A was utilized (Kaplan, *J. Immunol,* 151:628, 1993). 5000 colonies were grown overnight at 35° C. on LB agar plates containing 30 μg/ml ampicillin and 10 μg/ml tetracycline. 15–20 colonies were pooled and transferred to 10 ml Terrific Broth (containing the same antibiotics) and grown overnight. Plasmid DNA from each pool was then isolated by published procedures (Maniatis, Section 1.28), treated with RNAase (10 μg/ml), and stored in TE buffer.

The plasmid DNAs were transfected into COS-7(ATCC) cells using Lipofectamine (GIBCO/BRL, Gaithersburg, Md.). 1 μg of each plasmid pool was mixed with 3 μl Lipofectamine in 100 μl Opti-MEM medium (GIBCO/BRL), incubated for 30 min, and then placed on COS-7 cells that had been grown to 80–90% confluence in 12-well plates and rinsed in serum-free medium. The cells and DNA were incubated for 5 hrs at 37° C. with Dulbecco's MEM in the absence of serum and antibiotics, and then supplemented with the same medium containing 10% fetal calf serum and incubated overnight at 37° C. The next day, the medium was replaced with Dulbecco's MEM containing 10% fetal calf serum, penicillin, and streptomycin. After an additional 24 hrs of incubation, the medium was collected and stored at −20° C.

The cell supernatants were tested for IL-15 activity as described in Example 2 below. Five pools with the highest stimulation indices (1.6 to 2.1) exhibited levels of activity that were greater than 2 standard deviations from the mean of the remaining 278 pools. Three of the five pools remained positive in a second screen, and were subdivided into pools of 6. Plasmid DNA extracted from the secondary pools was used to transfect COS-7 cells and the supernatants were tested for IL-2-like activity. As described below in Example 2, three positive pools were identified and subdivided to yield individual clones; from each pool at least one positive clone was isolated.

The complete cDNA inserts of all three positive clones were sequenced using the automated Applied Biosystems Model 373A sequencing system. The flanking T7 and SP6 primers contained in the pcDNA1 vector were used to prime the sequencing reaction. Two of the clones, B2.16.2 and M2.12.1, were identical and coded for the cDNA sequence shown in FIG. 1. Clone F19.84 was similar to those two clones, but was missing the 20 nt at its 5' end (i.e., starting at the first ATG of the coding region) and contained a poly T tail of at least 100 nt at its 3' end.

The entire 747 nt sequence (FIG. 1, SEQ ID NO:1) was analyzed using a BLAST search (which accesses all of the major international nucleotide data banks). No significant homology was detected with any other known sequence. The sequence was also analyzed using the MacVector software program (MacVector 4.0; International Biotechnologies, Inc., New Haven, Conn.) on a Mac IIci computer. This analysis revealed an open reading frame flanked at its 5' end by a Kozak consensus sequence for translation initiation. The predicted amino acid sequence of this open reading frame is shown in FIG. 2 (Seq ID NO:2). This amino acid sequence was analyzed using a BLASTP search (which accesses all of the major international protein data banks) revealing significant homology with monkey and human precursor IL-15.

The predicted amino acid sequence of chicken IL-15 consists of a 143 amino acid polypeptide having a predicted molecular weight of 16,305 and an isoelectric point of 6.37. Based on the hydrophobicity of its amino terminal end and by comparison with known signal peptide cleavage sites (von Heijne, *Nucleic Acids Res.* 14:4683, 1986) it is predicted that cleavage between glycine-22 and alanine-23 results in the removal of an aminoterminal leader sequence of about 22 amino acids (secretion signal peptide) from the primary product to produce mature IL-15.

The predicted mature IL-15 sequence of 121 amino acids has a predicted molecular weight of 13,971, an isoelectric point of 6.57, and a possible N-linked glycosylation site (at asparagine 110 of FIG. 2). Comparisons between the predicted amino acid sequences of IL-15 from monkey, human, mouse and chicken and analysis of the tertiary structure of monkey IL-15 (Grabstein, *Science,* 264:965, 1994) suggest that four cysteines in chicken IL-15 (positions 63, 70, 116 and 119 of precursor IL-15, FIG. 2) are conserved and form intrachain disulfide bonds.

EXAMPLE 2

Bioactivity Assay for Chicken IL-15

Bioactivity assays for IL-15 are performed as follows: Concanavalin A (ConA)-activated splenic T cells are prepared by incubating chicken spleen cells (10^7 cells/ml) with Con A (10 μg/ml) (Sigma Chemical Co., St. Louis, Mo.) in RPMI 1640 medium (Sigma) containing 2 mg/ml BSA, antibiotics and glutamine at 40° C. for 24 hrs. The medium is then replaced with Iscoves' medium (Sigma) containing 2% normal chicken serum (Sigma) and 0.05M alpha-methyl pyrannoside (Sigma) for an additional 2–4 days, diluting the cells in additional medium as needed. Blast cells are purified from this mixture by gently layering them on a Histopaque density gradient (Sigma) and centrifuging them according to the manufacturer's instructions. The cells are then washed three times and finally resuspended in assay medium (Iscoves' containing 2% normal chicken serum (Sigma)).

For the assay, 2×10^4 blast cells are placed in roundbottom 96 well plates in assay medium containing IL-15 (such as, e.g., dilutions of supernatant from transfected COS-7 cells) or appropriate controls. After overnight incubation at 40° C.,

the cells are pulsed for 6 hrs with ^{3}H-thymidine (0.5 μCi) (New England Nuclear, Boston, Mass.)+fluorodeoxyuridine (10^{-6}M) (Sigma). The cells are then harvested on glass fiber filters (Whatman, Clifton, N.J.), and the radioactivity is measured in a liquid scintillation counter. IL-15 is expressed as a stimulation index, which is the radioactivity in experimental samples—the radioactivity in controls (non-transfected COS-7 supernatants). A typical result is shown in Table 1.

TABLE 1

SOURCE OF PLASMID DNA		Stimulation indices			
DNA	Designation	1/10 dil[a]	1/10 dil[b]	1/33 dil[b]	1/100 dil[b]
PRIMARY POOLS	A19	1.6	1.9	1.3	1.2
	B2	2.1	4.2	2.3	1.7
	E7	1.8	1.7	1.5	0.9
	F19	1.8	3.5	2.0	1.2
	M2	1.7	3.2	1.9	1.3
	Ave. of 278 ± SD	1.1 ± 0.1			
	Ave. of 3 Neg. pools		1.4	1.3	1.1
SECONDARY POOLS	A19.7		0.7	1.9	
	B2.16		6.0	3.5	
	F19.8		9.8	3.4	
	M2.12		3.2	2.2	
INDIVIDUAL CLONES	B2.16.2		6.6	3.3	2.7
	F19.8.4		7.5	4.0	3.0
	M2.12.1		7.2	3.9	3.6

[a]First screening at 1/10 dil.
[b]A repeat transfection using 5 positive and 3 negative primary pools

EXAMPLE 3

Expression and Purification of IL-15

To obtain high-level expression of chicken IL-15 in mammalian cells, the pSFV1 eukaryotic expression vector (which includes the Semliki Forest Virus replicon) is used (GIBCO/BRL, Gaithersburg, Md.). Use of this vector allows for signal peptide cleavage, glycosylation, and secretion of mature active protein. In one embodiment, the recombinant vector encodes an additional six histidine residues at the carboxyterminus of the native IL-15 sequence, allowing the efficient single step purification of the secreted protein on a nickel column (Novagen, Madison, Wis.).

Primers were constructed that include 5' and 3' sequences flanking the coding region of IL-15 cDNA. The 3' primer also includes nucleotides coding for 6 histidines. These primers were used in polymerase chain reaction (PCR), using as a template the entire IL-2 cDNA contained within the pcDNA1 plasmid. The resulting amplified cDNA, including the histidine-coding sequences, was ligated into the pSFV1 plasmid (GIBCO/BRL). The plasmid was obtained by transforming DH5 *E. coli* (GIBCO/BRL) and selecting transformants on agar plates and broth containing ampicillin.

This plasmid is used as a template to produce mRNA in vitro, using manufacturer's protocols. The mRNA is transfected into BHK-21 cells by electroporation, using 10 μg RNA per 10^7 cells, after which the cells are incubated for 1–3 days. The cell supernatant is harvested and passed through a resin matrix (His-Bind resin; Novagen, Madison, Wis.) using a suitable buffer system (His-bind buffer kit; Novagen). Up to 20 mg of tagged protein can be purified on a single 2.5 ml column. The IL-15 is eluted from the column with the elution buffer provided in the kit. It is estimated that BHK-21 cells growing in 50 ml medium synthesize about 25 mg total protein, with up to 5% comprising a recombinantly expressed and secreted protein. This corresponds to approximately 1.25 mg of cIL-15.

EXAMPLE 4

Use of Avian IL-15 in Vaccines

The following experiments are performed to evaluate the immune-enhancing activity of chicken IL-15 in chicken vaccines.

Chicken IL-15 cDNA is inserted into two viral vectors (derived from turkey herpesvirus and fowlpox virus, respectively) that are used for the expression of recombinant proteins in chickens (Morgan et al., *Avian Diseases,* 36:858, 1992; Yanagida et al., *J. Virol.,* 66:1402, 1992; Nazerian et al., *J. Virol,* 66:1409, 1992). These IL-15-modified live viral vectors are administered to newly hatched chicks simultaneously with the administration of various vaccines currently available. Six days later the chicks are challenged with the corresponding virulent viruses and observed for 8 weeks for the development of disease. The incidence of disease in these chicks is compared with controls that do not receive the IL-15-modified live viral vectors. A sample protocol (including expected results) is shown in Table 2.

TABLE 2

Group #	Treatment on day 1	Challenge at day 6	% expected with disease
1	none	none	0
2	none	virulent Marek's	>80%
3	HVT (not modified)	virulent Marek's	20%
4	HVT-IL-15[a]	virulent Marek's	0 to 10%
5	HVT (not modified) + HVT-IL-15	virulent Marek's	0 to 10%
6	none	virulent NDV	>80%
7	HVT-IF-15	virulent NDV	30% to >50%
8	NDV vaccine	virulent NDV	20%
9	NDV vaccine + HVT-IL-15	virulent NDV	0 to 10%

[a]herpesvirus of turkeys expressing IL-15

In an alternative procedure, newly hatched chicks are injected intramuscularly with 100 μg of a plasmid containing cDNA for chicken IL-15, using the methods described in Ulmer, *J. B. Science,* 259:1745–1749, 1993. These chicks, and control chicks receiving a control vector lacking IL-15 cDNA, are vaccinated on day 2 with chicken vaccines and

then challenged on day 7 with the corresponding virulent viruses. They are observed for 8 weeks for signs of disease. It is expected that chicks injected with the pcDNA1 vector containing IL-15 cDNA will exhibit a reduced incidence of disease relative to controls.

Finally, IL-15 protein purified by the procedure described in Example 3 is administered intramuscularly to chicks at hatching, followed by a single daily administration on each of the following four days. Chicks are divided into three groups, receiving 0.01, 0.1 or 1.0 μg per injection per day. A fourth group receives placebo injections. At hatching all chicks are vaccinated with chicken vaccines and then challenged on day 7 with the corresponding virulent viruses. They are then observed for 8 weeks for signs of disease. It is expected that chicks injected with IL-15 will exhibit a reduced incidence of disease relative to controls.

EXAMPLE 5

Use of Avian IL-15 in Growth Promotion

Mammalian IL-15 stimulates muscle growth (Quinn, L. S., *Endocrin.*, 136:3669, 1995) and semi-pure chicken IL-2 stimulates chicken body weight and increases feed conversion (U.S. Pat. No. 5,028,421). To evaluate the growth-promoting activity of avian IL-15, the methods described in Example 4 above may be used to administer IL-15 cDNA in a viral or plasmid vectors recombinant IL15 protein. Experimental and control chicks are monitored for weight gain and feed conversion for a period of six weeks. It is expected that one or more of these protocols will enhance chicken growth over controls.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 2

<210> SEQ ID NO 1
<211> LENGTH: 747
<212> TYPE: DNA
<213> ORGANISM: Gallus domesticus

<400> SEQUENCE: 1

cagataactg ggacactgcc atgatgtgca aagtactgat ctttggctgt atttcggtag 60

caacgctaat gactacagct tatggagcat ctctatcatc agcaaaaagg aaacctcttc 120

aaacattaat aaaggattta gaaatattgg aaaatatcaa gaacaagatt catctcgagc 180

tctacacacc aactgagacc caggagtgca cccagcaaac tctgcagtgt tacctgggag 240

aagtggttac tctgaagaaa gaaactgaag atgacactga aattaaagaa gaatttgtaa 300

ctgctattca aaatatcgaa aagaacctca agagtcttac gggtctaaat cacaccggaa 360

gtgaatgcaa gatctgtgaa gctaacaaca agaaaaaatt tcctgatttt ctccatgaac 420

tgaccaactt tgtgagatat ctgcaaaaat aagcaactaa tcatttttat tttactgcta 480

tgttatttat ttaattattt aattacagat aatttatata ttttatcccg tggctaacta 540

atctgctgtc cattctggga ccactgtatg ctcttagtct gggtgatatg acgtctgttc 600

taagatcata tttgatcctt tctgtaacct acgggctcaa aatgtacgtt ggaaaactga 660

ttgattctca ctttgtcggt aaagtgatat gtgtttactg aaagaatttt taaaagtcac 720

ttctagatga catttaataa atttcag 747

<210> SEQ ID NO 2
<211> LENGTH: 143
<212> TYPE: PRT
<213> ORGANISM: Gallus domesticus

<400> SEQUENCE: 2

Met Met Cys Lys Val Leu Ile Phe Gly Cys Ile Ser Val Ala Thr Leu
1 5 10 15

Met Thr Thr Ala Tyr Gly Ala Ser Leu Ser Ser Ala Lys Arg Lys Pro
20 25 30

Leu Gln Thr Leu Ile Lys Asp Leu Glu Ile Leu Glu Asn Ile Lys Asn
35 40 45

Lys Ile His Leu Glu Leu Tyr Thr Pro Thr Glu Thr Gln Glu Cys Thr
50 55 60

-continued

```
Gln Gln Thr Leu Gln Cys Tyr Leu Gly Glu Val Val Thr Leu Lys Lys
65                  70                  75                  80

Glu Thr Glu Asp Asp Thr Glu Ile Lys Glu Glu Phe Val Thr Ala Ile
                85                  90                  95

Gln Asn Ile Glu Lys Asn Leu Lys Ser Leu Thr Gly Leu Asn His Thr
            100                 105                 110

Gly Ser Glu Cys Lys Ile Cys Glu Ala Asn Asn Lys Lys Lys Phe Pro
        115                 120                 125

Asp Phe Leu His Glu Leu Thr Asn Phe Val Arg Tyr Leu Gln Lys
    130                 135                 140
```

What is claimed is:

1. An isolated nucleic acid which:

(a) comprises a nucleic acid sequence having at least 70% sequence homology, determined by a BLAST algorithm, to the sequence set forth in nucleotides 87–449 of SEQ ID NO:1; and

(b) encodes a polypeptide capable of stimulating thymidine incorporation in mitogen activated avian T-cells.

2. The complement of a nucleic acid according to claim **1**.

3. An isolated nucleic acid according to claim **1** or **2**, which nucleic acid is an avian nucleic acid isolated from chicken.

4. A vector construct comprising the nucleic acid of claim **1**.

5. The vector construct according to claim **4**, in which said nucleic acid is operatively associated with a promoter element capable of expressing the nucleic acid in a host cell.

6. The vector construct according to claim **4**, in which the construct is a recombinant virus.

7. The vector construct according to claim **6**, in which the recombinant virus is a turkey herpes virus or a fowl pox virus.

8. An isolated nucleic acid which:

(a) hybridizes to the full length of a nucleic acid having the complementary sequence of nucleotides 87–449 in SEQ ID NO:1 under conditions comprising (i) hybridization in 6×SSC and 0.5% SDS, and (ii) washing at 68° C. in 0.1×SSC and 0.5% SDS; and

(b) encodes a polypeptide capable of stimulating thymidine incorporation in mitogen activated avian T-cells.

9. The complement of a nucleic acid according to claim **8**.

10. An isolated nucleic acid according to claim **8** or **9**, which nucleic acid is an avian nucleic acid isolated from chicken.

11. An isolated nucleic acid which:

(a) hybridizes to the full length of a nucleic acid having the complementary sequence of nucleotides 87–449 in SEQ ID NO:1 under conditions comprising (i) hybridization in 6×SSC and 0.5% SDS, and (ii) washing at room temperature in 2×SSC and 0.5% SDS; and

(b) encodes a polypeptide capable of stimulating thymidine incorporation in mitogen activated avian T-cells.

12. The complement of a nucleic acid according to claim **11**.

13. An isolated nucleic acid according to claim **11** or **12**, which nucleic acid is an avian nucleic acid isolated from chicken.

14. A vector construct comprising the nucleic acid of claim **8** or **11**.

15. A The vector construct according to claim **14**, in which said nucleic acid is operatively associated with a promoter element capable of expressing the nucleic acid in a host cell.

16. The vector construct according to claim **14**, in which the construct is a recombinant virus.

17. The vector construct according to claim **16**, in which the recombinant virus is a turkey herpes virus or a fowl pox virus.

18. An isolated nucleic acid having an open reading frame that encodes a polypeptide comprising the sequence of amino acid residues 23–143 set forth in SEQ ID NO:2 (FIG. **2**).

19. The complement of a nucleic acid according to claim **18**.

20. An isolated nucleic acid according to claim **18** or **19**, which nucleic acid is an avian nucleic acid isolated from chicken.

21. An isolated nucleic acid according to claim **18**, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO:2 (FIG. **2**).

22. The complement of a nucleic acid according to claim **21**.

23. An isolated nucleic acid according to claim **21** or **22** which nucleic acid is an avian nucleic acid isolated from chicken.

24. A vector construct comprising the nucleic acid of claim **18** or **21**.

25. The vector construct according to claim **24**, in which said nucleic acid is operatively associated with a promoter element capable of expressing the nucleic acid in a host cell.

26. The vector construct according to claim **24**, in which the construct is a recombinant virus.

27. The vector construct according to claim **26**, in which the recombinant virus is a turkey herpes virus or a fowl pox virus.

* * * * *